

Methodology for the Assessment of Spatial Economic Impacts of Transport Projects and Policies

**Johannes Bröcker
d'Artis Kancs
Carsten Schürmann
Michael Wegener**

September 2001

Abstract

In this study we develop an integrated econometric and CGE modelling framework for transport projects and transport policies at the European level by integrating network, regional economic and macro-economic impacts. The paper presents the formal structure of the integrated econometric and CGE modelling framework, explains the calibration and applies it to the policy evaluation. The effects of infrastructure investments are modelled by simulating trade cost changes in a comparative static analysis, using estimates of trade cost changes due to new infrastructure links, obtained from a transport network model. By performing a systematic and quantitative analysis of the spatial, network and socio-economic impacts of transport investments and policy and carrying out scenario simulations we improve the understanding of the impact of transportation policies on short- and long-term spatial development in the EU.

Keywords: Transport Policy, Impact analysis, Spatial CGE Model, Trade cost.

JEL classification: C68, D58, O41, R13

Contents

	Page
1 Introduction	4
2 The Extended SASI Model	7
2.1 Theoretical Approaches	7
2.2 Model Overview	9
2.3 Model Extensions	11
2.4 Submodels	13
2.4.1 European Developments	13
2.4.2 Regional Accessibility	14
2.4.3 Regional GDP	16
2.4.4 Regional Employment	18
2.4.5 Regional Population	19
2.4.6 Regional Labour Force	21
2.4.7 Socio-economic Indicators	22
2.4.8 Cohesion Indicators	23
2.5 Other Model Output	24
2.6 Data Requirements	26
2.6.1 Simulation Data	27
2.6.2 Calibration/Validation Data	28
3 The New CGEurope Model	29
3.1 Modelling the Spatial Dimension of Transport Benefits	30
3.2 Extending CGEurope	30
3.3 Non-Formal Description of the New CGEurope	32
3.4 The Mathematical Structure of the New CGEurope Model	35
3.4.1 The Equations of the Equilibrium	35
3.4.2 Household Behaviour and Welfare Measurement	38
3.4.3 Transport Costs	38
3.4.4 Calibration	40
3.4.5 Model results to be Used in Project Assessment	43
3.5 The CGEurope Model Data Requirements	44
3.5.1 Regional Data	45
3.5.2 National Accounting Data	47
3.5.3 Economic Sectors	48
3.5.4 Transport Costs	50
3.5.5 Passenger Transport Flows	50
4 Regions and Networks	51
4.1 System of Regions	51
4.2 Trans-European Networks	51
5 Possible Scenarios	62
5.1 Transport Infrastructure Scenarios	63
5.2 Socio-Economic Scenarios	63
5.3 Scenario Applications	64
6 Conclusions	65
7 References	67

1 Introduction

The goal of the present study is to improve the understanding of the impact of transportation policies on short- and long-term spatial development in the EU by developing a *unified assessment framework* for transport project and transport policies at the European level integrating network, regional economic and macro-economic impacts.

At the heart of the present study lie the following activities:

- improvement of existing assessment frameworks by ensuring that direct and indirect impacts are clearly distinguished within the appraisal, and that the incidence of benefits and costs, and sources of additionality and/or double-counting are transparent.
- a systematic and quantitative analysis of the network, spatial and socio-economic impacts of transport investments and policies by refining existing EU-level models and carrying out scenario simulations.
- building up and maintaining a discussion platform for the cluster in order to facilitate interactions between Subtask 5, the other subtasks of the cluster and the scientific community.
- building guidelines and recommendations for project analysis of transport investments and policies and for the development of supporting tools and databases, in order to improve the applicability of the outputs of the project in policy analysis.

The main innovations of the present study lie in the field of the methodological research on project assessment, its further development in order to allow applications at the EU level and the proposed modelling approach, which has not yet been applied in this conjunction before. Based upon this structure, at the end of the project an effective tool to evaluate in a coherent way the aspects of sustainability, cohesion, environmental-friendliness and efficiency of transport policies and projects will exist.

Structure of the study

The study is structured as follows:

The following two sections describe the extended SASI model (Section 2) and the new CGEurope model (Section 3) in a similar fashion, starting with a general model overview describing briefly the overall structure and functionality of the models, followed by a description of the model extensions and concluded by data requirements as a basis for the establishment of the common spatial database.

Section 4 describes the system of regions agreed upon, the trans-European transport networks established, and the sectoral categorisation to be implemented. Taking these three aspects together, they define the general framework of the models and the database to be set up.

Section 5 gives first suggestions on a general framework for scenario applications to be tested based on the structure and capabilities of the two models, notwithstanding the work to be done in Task 2.3 *Definition of Transport Policy Scenarios*. The scenario framework proposed here is mainly based on scenario applications done within the SASI project, but suggests also additional scenarios not tested in SASI with respect to the model extensions.

Section 6 concludes the Report in summarising the main findings and giving a preview of future work.

2 The Extended SASI Model

The important role of transport infrastructure for regional development is one of the fundamental principles of regional economics. In its most simplified form it implies that regions with better access to the locations of input materials and markets will, *ceteris paribus*, be more productive, more competitive and hence more successful than more remote and isolated regions (Jochimsen, 1966). However, the relationship between transport infrastructure and economic development seems to be more complex than this simple model. There are successful regions in the European core confirming the theoretical expectation that location matters. However, there are also centrally located regions suffering from industrial decline and high unemployment. On the other side of the spectrum the poorest regions, as theory would predict, are at the periphery, but there are also prosperous peripheral regions such as the Scandinavian countries. To make things even more difficult, some of the economically fastest growing regions are among the most peripheral ones.

So it is not surprising that it has been difficult to empirically verify the impact of transport infrastructure on regional development. There seems to be a clear positive correlation between transport infrastructure endowment or the location in interregional networks and the *levels* of economic indicators such as GDP per capita (e.g. Biehl, 1986; 1991; Keeble et al., 1982; 1988). However, this correlation may merely reflect historical agglomeration processes rather than causal relationships still effective today (cf. Bröcker and Peschel, 1988). Attempts to explain *changes* in economic indicators, i.e. economic growth and decline, by transport investment have been much less successful. The reason for this failure may be that in countries with an already highly developed transport infrastructure further transport network improvements bring only marginal benefits. The conclusion is that transport improvements have strong impacts on regional development only where they result in removing a *bottleneck* (Blum, 1982; Biehl, 1986; 1991).

2.1 Theoretical Approaches

There exists a broad spectrum of theoretical approaches to explain the impacts of transport infrastructure investments on regional socio-economic development. Originating from different scientific disciplines and intellectual traditions, these approaches presently coexist, even though they are partially in contradiction (cf. Linnecker, 1997):

- *National growth approaches model* multiplier effects of public investment in which public investment has either positive or negative (crowding-out) influence on private investment, here the effects of transport infrastructure investment on private investment and productivity. In general only national economies are studied and regional effects are ignored. Pioneered by Aschauer (1989; 1993) such studies use time-series analyses and growth model structures to link public infrastructure expenditures to movements in private sector productivity. An increase in public investment raises the marginal product of private capital and provides an incentive for a higher rate of private capital accumulation and labour productivity growth. Critics of these approaches argue that there may be better infrastructure strategies than new construction and that policy measures aimed at increasing private investment directly rather than via public investment will have greater impact on national competitiveness.

- *Regional growth approaches* rest on the neo-classical growth model which states that regional growth in GDP per capita is a function of regional endowment factors including public capital such as transport infrastructure, and that, based on the assumption of diminishing returns to capital, regions with similar factors should experience converging per-capita incomes over time. The suggestion is that, as long as transport infrastructure is unevenly distributed among regions, transport infrastructure investments in regions with poor infrastructure endowment will accelerate the convergence process, whereas once the level of infrastructure provision becomes uniform across regions, they cease to be important. Critics of regional growth models built on the central assumption of diminishing returns to capital argue that they cannot distinguish between this and other possible mechanisms generating convergence such as migration of labour from poor to rich regions or technological flows from rich to poor regions.
- *Production function approaches* model economic activity in a region as a function of production factors. The classical production factors are capital, labour and land. In modern production function approaches infrastructure is added as a public input used by firms within the region (Jochimsen, 1966; Buhr, 1975). The assumption behind this expanded production function is that regions with higher levels of infrastructure provision will have higher output levels and that in regions with cheap and abundant transport infrastructure more transport-intensive goods will be produced. The main problems of regional production functions are that their econometric estimation tends to confound rather than clarify the complex causal relationships and substitution effects between production factors. This holds equally for production function approaches including measures of regional transport infrastructure endowment. In addition the latter suffer from the fact that they disregard the network quality of transport infrastructure, i.e. treat a kilometre of motorway or railway the same everywhere, irrespective of where they lead to.
- *Accessibility approaches* attempt to respond to the latter criticism by substituting more complex accessibility indicators for the simple infrastructure endowment in the regional production function. Accessibility indicators can be any of the indicators discussed in Schürmann et al. (1997), but in most cases are some form of population or economic potential. In that respect they are the operationalisation of the concept of 'economic potential' which is based on the assumption that regions with better access to markets have a higher probability of being economically successful. Pioneering examples of empirical potential studies for Europe are Keeble et al. (1982; 1988). Today approaches relying only on accessibility or potential measures have been replaced by the hybrid approaches where accessibility is but one of several explanatory factors of regional economic growth. Also the accessibility indicators used have become much more diversified by type, industry and mode (see Schürmann et al., 1997). The SASI model is a model of this type incorporating accessibility as one explanatory variable among other explanatory factors.
- *Regional input-output approaches* model interregional and inter-industry linkages using the Leontief (1966) multiregional input-output framework. These models estimate inter-industry/interregional trade flows as a function of transport cost and a fixed matrix of technical inter-industry input-output coefficients. Final demand in each region is exogenous. Regional supply, however, is elastic, so the models can be used to forecast regional economic development. One recent example of an operational multiregional input-output model is the MEPLAN model (Marcial Echenique & Partners Ltd., 1998).

- *Trade integration approaches* model interregional trade flows as a function of interregional transport and regional product prices. Peschel (1981) and Bröcker and Peschel (1988) estimated a trade model for several European countries as a doubly-constrained spatial interaction model with fixed supply and demand in each region in order to assess the impact of the economic integration of Europe in terms of reduced tariff barriers and border delays between European countries. Their model could have been used to forecast the impacts of transport infrastructure improvements on interregional trade flows. If the origin constraint of fixed regional supply were relaxed, the model could have been used also for predicting regional economic development. Krugman (1991) and Krugman and Venables (1995) extended this simple model of trade flows by the introduction of economies of scale and labour mobility. The CGEurope presented in Section 3 is a model of this type.

In this section, the extended SASI model will be presented.

2.2 Model Overview

The SASI model is a recursive simulation model of socio-economic development of regions in Europe subject to exogenous assumptions about the economic and demographic development of the European Union as a whole and transport infrastructure investments and transport system improvements, in particular of the trans-European transport networks (TETN).

The main concept of the SASI model is to explain locational structures and locational change in Europe in combined time-series/cross-section regressions, with accessibility indicators being a subset of a range of explanatory variables. Accessibility is measured by spatially disaggregate accessibility indicators which take into account that accessibility within a region is not homogenous but rapidly decreases with increasing distance from the nodes of the networks. The focus of the regression approach is on long-term spatial distributional effects of transport policies. Factors of production including labour, capital and knowledge are considered as mobile in the long run, and the model incorporates determinants of the redistribution of factor stocks and population. The model is therefore suitable to check whether long-run tendencies in spatial development coincide with development objectives discussed above. Its application is restricted, however, in other respects: The model generates distributive, not generative effects of transport cost reductions, and it does not produce regional welfare assessments fitting into the framework of cost-benefit analysis.

The SASI model differs from other approaches to model the impacts of transport on regional development by modelling not only production (the demand side of regional labour markets) but also population (the supply side of regional labour markets), which makes it possible to model regional unemployment. A second distinct feature is its dynamic network database based on a 'strategic' subset of highly detailed pan-European road, rail and air networks including major historical network changes as far back as 1981 and forecasting expected network changes according to the most recent EU documents on the future evolution of the trans-European transport networks.

The SASI model has six forecasting submodels: *European Developments*, *Regional Accessibility*, *Regional GDP*, *Regional Employment*, *Regional Population* and *Regional Labour Force*. A seventh submodel calculates *Socio-Economic Indicators* with respect to efficiency and equity. Figure 2.1 visualises the interactions between these submodels.

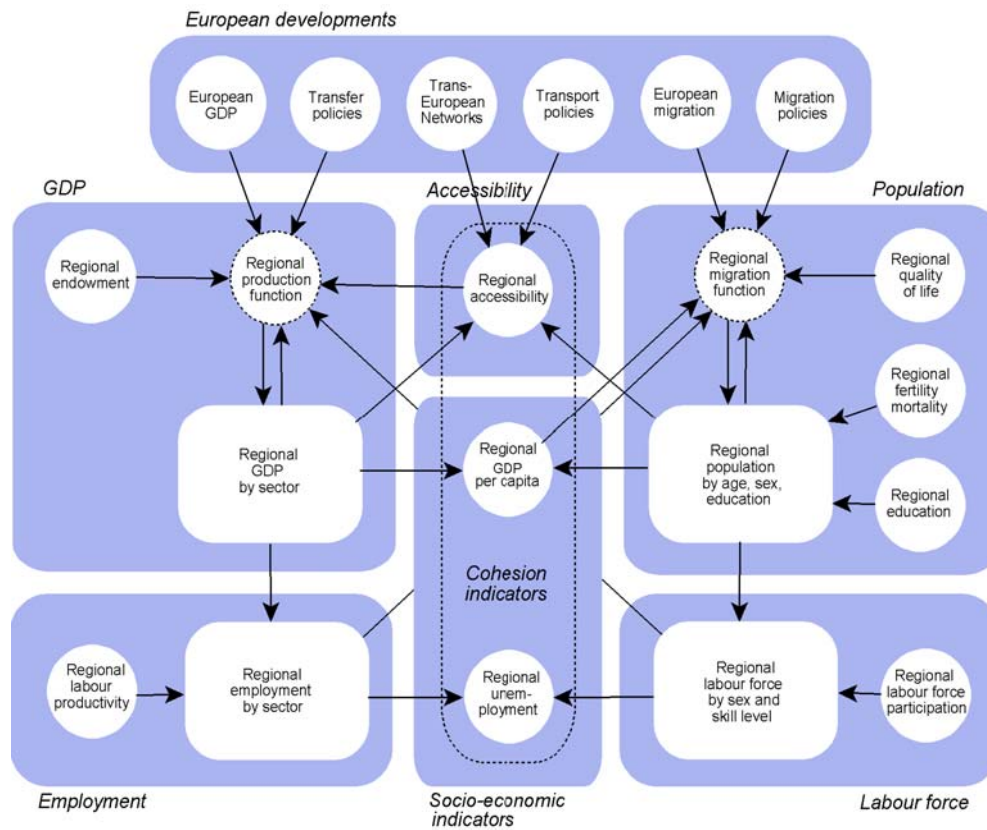


Figure 1. The SASI model

The *spatial* dimension of the model is established by the subdivision of the European Union and the 12 candidate countries in eastern Europe in 1,245 regions and by connecting these regions by road, rail and air networks (see Section 4). For each region the model forecasts the development of accessibility, GDP per capita and unemployment. In addition cohesion indicators expressing the impact of transport infrastructure investments and transport system improvements on the convergence (or divergence) of socio-economic development in the regions of the European Union are calculated.

The *temporal* dimension of the model is established by dividing time into periods of one year duration. By modelling relatively short time periods both short- and long-term lagged impacts can be taken into account. In each simulation year the seven submodels of the SASI model are processed in a recursive way, i.e. sequentially one after another. This implies that within one simulation period no equilibrium between model variables is established; in other words, all endogenous effects in the model are lagged by one or more years.

The mathematical specification of the original SASI model is contained in EUNET/SASI Report 8 (Wegener and Böckmann, 1998). The implementation of the original SASI model, i.e. the application of empirical data to it and the estimation and calibration of its parameters, was described in EUNET/SASI Report 11 (Füürst et al., 1999). The software system of the original SASI model was described in EUNET/SASI Report 13 (Wegener et al., 2000). The results of the demonstration scenario simulations with the original SASI model were presented in EUNET/SASI Report D15 (Füürst et al., 2000).

2.3 Model Extensions

This section presents the extended SASI model. The presentation follows the description of the original SASI model in Wegener and Böckmann (1998) and Fürst et al. (1999) but focuses on the modifications of and extensions to the original model specification. In the present study, the SASI model will be updated and extended in several dimensions relating to *model theory*, *model data* and *model technique*. Before the extended model will be presented in detail, these model extensions are summarised.

Model Theory

New ideas from growth theory as well as new evidence on firm location will be reviewed and transformed into operational indicators of locational advantage and disadvantage and incorporated into the econometric approach. The following changes are intended:

- *Rates v. levels*. The traditional production function approach relates the *level* of output to the *level* of infrastructure. New growth theory suggests that a link might also exist between the level of infrastructure and the *rate* of growth, because good accessibility means good access to diversity making research and development more productive. It will be examined whether this effect can be incorporated into the model functions by exploring the feasibility of forecasting *rates* of change of regional economic development rather than the *levels* of regional production,
- *Productivity*. The feasibility of forecasting regional sectoral labour productivity endogenously as a function of accessibility and other variables instead of using exogenous productivity forecasts will be explored.
- *Accessibility*. The possibility to explicitly consider wage levels and/or production costs of potential suppliers in other regions in the accessibility submodel will be examined. It is expected that this would enhance the contribution of the accessibility indicators to the explanation of regional economic development in the regional production functions.
- *Migration*. It is planned to forecast migration *flows* as a function of regional unemployment and other indicators expressing the attractiveness of the region as a place of employment and a place to live instead of the present *net* migration. It is expected that this will improve the explanatory power of the migration model in the Population submodel.

In addition, efforts will be made to make the model more policy-relevant:

- *Policies*. The model will be made more responsive to non-transport policies, such as regional economic policies or immigration policies, and to a broader range of transport policies, such as policies addressing intermodality and congestion. This work will build on the definition of transport policy scenarios in Task 2.3 (see Section 5).
- *Cohesion indicators*. The cohesion indicators used for assessing the impacts of transport policies will be expanded and critically assessed with respect to their possible implicit bias towards convergence and divergence. One of the findings of the SASI project was that the choice of cohesion indicator, i.e. whether relative or absolute differences are calculated, is critical for whether transport infrastructure projects have a cohesion effect or contribute to spatial polarisation.

Model Data

Work is underway to implement the common spatial model database to be used by both the CGEurope and the extended SASI model. This incorporates the following steps:

- *Disaggregation.* The existing SASI regional model database is presently being disaggregated from 201 NUTS-2 regions to 1,083 NUTS-3 regions in the present 15 member states of the European Union (see Section 4) and to include six economic sectors instead of the previous three. In order to be compatible, the following six economic sectors will be considered (see Section 3):
 - Manufactured products
 - Market services
 - Agriculture, forestry and fishery products
 - Fuel and power products
 - Building and construction
 - Non-market services
- *Updating.* The resulting 1,083-region database will be updated to include more recent data.
- *Extension.* The database for calibration/validation will be extended by additional variables, such as labour productivity and wage levels and/or production costs by sector.
- *Candidate countries.* A similar model database for the 162 regions in the 12 candidate countries (see Section 4) will be established.
- *Transport networks.* The road, rail and air transport networks to be used by the two models are being refining, extended and updated to include the 12 candidate countries and the related extensions of the trans-European networks, to connect the new high-resolution system of regions and to incorporate expected network changes after 2016 until 2021.

Model Technique

One of the results of the SASI project was that the state of the art of calibrating and validating dynamic models of the kind of SASI *over time* is poorly developed. Efforts will therefore be made to calibrate and validate the extended SAS model with time-series data of regions and countries, also with respect to model variables not considered as output indicators.

In addition, work has started on the extension of the SASI model software system in three respects:

- The model dimensions were extended to incorporate the new system of regions with up to 1,500 regions.
- The model software was ported to a software development environment with full Windows integration with multiple windows, dialog boxes and pull-down menus.
- The graphical user interface of the model was enhanced by visual output in the form of on-line time-series plots, choropleth maps and 3D representations of spatial distributions, as well as offline comparison between simulated scenarios.

2.4 Submodels

In this section the specification of model variables and parameters of the seven submodels of the SASI model is presented in detail. The description of each submodel starts with the specification of the submodels in the original SASI model in Wegener and Böckmann (1998) and Fürst et al. (1999) and then the modifications in the model specification which are planned are pointed out and explained.

2.4.1 European Developments

The *European Developments* submodel is not a 'submodel' in the narrow sense because it simply prepares exogenous assumptions about the wider economic and policy framework of the simulations and makes sure that external developments and trends are considered.

For each simulation period the simulation model requires the following assumptions about European developments:

- (1) *Assumptions about the performance of the European economy as a whole.* The performance of the European economy is represented by observed values of sectoral GDP for the European Union as a whole and for 26 non-EU countries (see Table 4.1 and Figure 4.1) for the years 1981 to 1997 and forecasts for the years 1998 to 2016. All GDP values are entered in Euro in prices of 2001.
- (2) *Assumptions about immigration and outmigration across Europe's borders.* European migration trends are represented by observed annual immigration and outmigration of the EU member states and the EU as a whole for the years 1981 to 1997 and of forecasts for the years 1998 to 2016.

These two groups of assumptions serve as constraints to ensure that the regional forecasts of economic development and population remain consistent with external developments not modelled. To keep the total economic development exogenous to the model means that the model is prevented from making forecasts about the general increase in production through transport infrastructure investments, although in principle its parameters are estimated in a way that makes it capable of doing that. Alternatively, it is possible to let the model determine the total level of annual GDP and to use the observed values of the period from 1981 to 1997 to validate these forecasts.

- (3) *Assumptions about transfer payments by the European Union via the Structural Funds and the Common Agricultural Policy or by national governments to support specific regions.* European and national transfer payments are taken into account by annual transfers (in Euro of 2001) received by the regions in the European Union during the period 1981 to 1997 and forecasts for the period 1998 to 2016. These data are provided only for those regions that actually received financial support in the past or are assumed to receive support in the future.
- (4) *Assumptions about immigration policies by European countries.* Given the expected rapid population growth and lack of economic opportunity in many origin countries, total European immigration will largely be influenced by policy decisions by national governments. These assumptions are reflected by upper limits for annual immigration from non-EU countries to the countries of the European Union for the years 1981 to 1997 and forecasts for the years 1998 to 2016.

The data for these four types of assumptions do not need to be provided for each year nor for time intervals of equal length as the model performs the required interpolations for the years in between.

- (5) *Assumptions about the development of trans-European transport networks (TETN).* The European road, rail and air networks are backcast for the period between 1981 and 1996 and, based on assumptions on the development of trans-European networks, forecast until the year 2016, both in five-year increments. The *base forecast* or *base scenario* is defined as the implementation of the fourteen trans-European transport network priority projects approved on the Essen summit (see Section 5).
- (6) *Assumptions about policy decisions on the trans-European networks.* A *policy scenario* is a time-sequenced investment programme for addition, upgrading or closure of links of the trans-European road, rail or air networks. Policy scenarios are specified by adding different subsets of the remaining TETN links such as all planned TETN road projects, all planned TETN rail projects or all planned TETN road and rail projects (see Section 5).

2.4.2 Regional Accessibility

This submodel calculates regional accessibility indicators expressing the locational advantage of each region with respect to relevant destinations in the region and in other regions as a function of the travel time or travel cost (or both) needed to reach these destinations by the strategic road, rail and air networks.

The method to calculate accessibility indicators used for the SASI model was described in Schürmann et al. (1997) and Fürst et al. (1999): The European territory was subdivided into some 70,000 raster cells of 10 kilometres width and population of 1995 and GDP of 1992 were disaggregated to raster cells assuming a negative-exponential density gradient around major cities. These distributions were used during the simulation as ancillary information to allocate population and GDP as predicted by the model for each region to the raster cells belonging to that region. Accessibility was calculated in each year for each region by using population and GDP in the 70,000 raster cells in Europe as destinations.

For the selection of accessibility indicators to be used in the model three, possibly conflicting, objectives were considered to be relevant: First, the accessibility indicators should contribute as much as possible to explaining regional economic development. Second, the accessibility indicators should be meaningful by itself as indicators of regional quality of life. Third, the accessibility indicators should be consistent with theories and empirical knowledge about human spatial perception and behaviour.

In the light of these objectives, three types of accessibility indicator were tested (see Schürmann et al., 1997):

- (1) *travel time or cost:* average travel time or cost from each region centroid to a predefined set of destinations,
- (2) *daily accessibility:* the total of population or GDP that can be reached from the region within a certain time or cost limit,

(3) *potential accessibility*: the total of destination activities, population or GDP, $W_j(t)$, in all 70,000 destination cells j in year t weighted by a negative exponential function of travel time or cost by mode m between centroid cell k and destination cells j :

$$A_{rm}(t) = \sum_j W_j(t) \exp[-\beta c_{kjm}(t)] \quad (2.1)$$

where $A_{rm}(t)$ is the accessibility of region r by mode m in year t , and is k the raster cell of the centroid of region r .

Of these, potential accessibility was adopted.

As modal impedance $c_{kjm}(t)$ rail timetable travel times and road travel times calculated from road-type specific travel speeds were used. In addition, following Bröcker (1984; 1996), political and cultural barriers were taken account of as time penalties added to the travel times:

$$c_{kjm} = c'_{kjm}(t) + e_{k'j'}(t) + \ell_{k'j'} + s_{k'j'}(t) \quad (2.2)$$

in which $c'_{kjm}(t)$ is the pure travel time between centroid cell k and destination cells j by mode m in year t and $e_{k'j'}(t)$, $\ell_{k'j'}$ and $s_{k'j'}$ are exogenous time penalties for political and cultural diversity in year t between the countries k' and j' to which cells k and j belong:

- $e_{k'j'}(t)$ is a *European integration factor* reflecting in which supranational structures the two countries are embedded, i.e. which political and economic relationship existed between them in year t ,
- $\ell_{k'j'}$ is a *language factor* describing the grade of similarity of the mother language(s) spoken in the two countries
- $s_{k'j'}$ is a *cultural similarity factor* reflecting how similar are cultural and historical experience of the two countries.

While the latter two factors were kept constant over the whole simulation, $e_{k'j'}(t)$ was reduced from year to year to account for the effect of European integration. For the specification of the three factors, see Fürst et al. (1999). The accessibility indicators used in the model were not standardised to the European average to show increases in accessibility over time.

Modal accessibility indicators were aggregated to one indicator expressing the combined effect of alternative modes by replacing the impedance term $c_{kjm}(t)$ by the composite or *logsum* impedance (Williams, 1977):

$$c_{kj}(t) = -\frac{1}{\lambda} \ln \sum_{m \in M_{kj}} \exp[-\lambda c_{kjm}(t)] \quad (2.3)$$

where M_{kj} is the set of modes available between raster cells k and j .

Two composite accessibility indicators were used: the logsum of road and rail accessibility to population (population potential) for the agricultural and industrial sectors, and the logsum of road, rail and air accessibility to GDP (economic potential) for the service sector.

In the extended SASI model travel time will be replaced by *generalised cost*, i.e. an aggregate measure of both time and cost of transport. Travel costs will be calculated from link-type spe-

cific cost parameters. Travel time will be converted to cost by appropriate assumptions about the value of time of travellers and drivers. The border penalties mentioned above, which were converted to time units in the original SASI model, will be converted to monetary cost equivalents in the extended SASI model.

In addition, it will be explored whether an extension of the accessibility concept taking account of differences between regions in the prices of commodities as proposed by Bröcker (1996) will bring a gain in explanatory power of accessibility indicators. Bröcker suggested to distinguish between supply potential

$$S_r(t) = \sum_s V_s(t) p_s^{-\sigma}(t) \exp[(1-\sigma)c_{sr}] \quad (2.4)$$

and demand potential

$$D_r(t) = \sum_s \frac{V_s(t)}{\sum_r V_r(t) p_r^{-\sigma}(t) \exp(-\sigma c_{sr})} \exp(-\sigma)c_{rs} \quad (2.5)$$

where V_s is output of region s , p_s is the price of commodities produced in s , c_{rs} and c_{sr} are generalised costs as explained above, and σ is the elasticity of substitution between commodities and regions. The rationale of the two potentials is that, unlike the economic potential commonly used, they measure only those parts of the supply or distribution markets in regions s that are relevant for producers in region r because of their competitive prices.

The feasibility of using supply and demand potentials instead of economic and population potential in the extended SASI model will depend on the possibility to find appropriate proxies for V_s and p_s . V_s may be approximated by GDP, and the elasticity of substitution parameter σ may be provided by the CGEurope model. However, the determination of product prices p_s of each year will be difficult as the SASI model does not have an explicit model of commodity markets. It will be explored whether commodity prices can be approximated by wage levels because wage differentials may be a primary factor of differences in commodity prices. Such a strategy would show the comparative advantage of peripheral low-wage regions, in particular in the candidate countries in eastern Europe.

2.4.3 Regional GDP

The GDP submodel is based on a quasi-production function incorporating accessibility as additional production factor. The economic output of a region is forecast separately for each economic sector (agriculture, manufacturing, services) in order to take different requirements for production by each sector into account. The regional production function predicts annual regional GDP:

$$Q_{ir}(t) = f[C_{ir}(t), L_{ir}(t), A_{ir}(t), S_r(t), R_{ir}] \quad (2.6)$$

where $Q_{ir}(t)$ is annual GDP of industrial sector i in region r in year t , $C_{ir}(t)$ is a vector of endowment factors relevant for industrial sector i in region r in year t , $L_{ir}(t)$ is labour relevant for industrial sector i in region r in year t , A_{ir} is a vector of accessibility indicators relevant for industrial sector i in region r in year t , $S_r(t)$ are annual transfers received by the region r in year t and R_{ir} is a region-specific residual taking account of factors not modelled (see below).

Note that, even though annual GDP is in fact a flow variable relating to a particular time interval (year), it is modelled like a stock variable.

Sectoral GDP, however, does not only contain the actual economic output of a region, but also includes different kinds of subsidies, which can apparently not be explained by the production conditions within a region. Consequently, sectoral GDP has to be reduced by transfer payments and their multiplier effects (see Fürst et al., 1999, Section 4.1.2). The production function then becomes

$$Q_{ir}(t) - S_{ir}(t) = f[C_{ir}(t), L_{ir}(t), A_{ir}(t), R_{ir}] \quad (2.7)$$

Assuming that the different production factors can be substituted by each other only to a certain degree, a multiplicative function which reflects a limitational relation between the factors was chosen. Since this kind of function introduces the coefficients as exponents of the explaining variables it is possible to interpret the coefficients as elasticities of production reflecting the importance of the different production factors for economic growth in a sector.

Due to different ways of production, each economic sector depends on different production conditions and factors. Therefore the three sectoral functions contain different explanatory variables. In spite of that the functions show the same basic structure. All of them use the following four types of explanatory variables: regional labour force, accessibility, economic structure and endowment factors. The operational specification of the regional production functions used in the SASI model is:

$$Q_{ir}(t) - S_{ir}(t) = L_{ir}(t)^\alpha A_{ir}(t)^\chi q_{ir}(t-1)^\beta X_{ir}(t)^\delta \varepsilon + R_{ir}(t) \quad (2.8)$$

where $Q_{ir}(t)$ is economic output (GDP) of sector i in region r in year t , $S_{ir}(t)$ are transfer payments in region r relevant for sector i in year t , $L_{ir}(t)$ is labour force in region r in year t , $A_{ir}(t)$ is accessibility of region r relevant for sector i in year t , $q_{ir}(t-1)$ is the economic structure of region r (sectoral share of sector i in year $t-1$), $X_{ir}(t)$ is an aggregate of endowment factors in region r relevant for sector i in year t , $R_{ir}(t)$ denotes regression residuals of the estimated GDP values of sector i in region r in year t and α , β , χ , δ , and ε are regression coefficients.

The economic structure variable was used as an explanatory variable on the grounds that the conditions for production in a certain sector depend heavily on the given sectoral structure, a fact which reflects historic developments and path dependencies that are not covered by any other indicator in the equation. Endowment factors are indicators measuring the suitability of the region for economic activity. Endowment factors include traditional location factors such as capital stock (i.e. production facilities) and intraregional transport infrastructure as well as 'soft' quality-of-life factors such as indicators describing the spatial organisation of the region, i.e. its settlement structure and internal transport system, or institutions of higher education, cultural facilities, good housing and a pleasant climate and environment (for the specification of the quality-of-life indicator, see Schürmann, 1999). In addition to endowment factors and accessibility indicators, monetary transfers to regions by the European Union such as assistance by the Structural Funds or the Common Agricultural Policy or by national governments are considered, as these account for a sizeable portion of the economic development of peripheral regions. Regional transfers per capita $S_r(t)$ are provided by the *European Developments* submodel (see Section 2.4.1). To take account of 'soft' factors not captured by the endowment and accessibility indicators of the model, a region- and sector-specific residual con-

stant R_{ir} is added to the GDP forecasts of each region r . R_{ir} is the difference between the GDP per capita predicted for region r in the base year 1981 and observed GDP per capita in r in 1981. R_{ir} is kept constant over all simulation periods.

The results of the regional GDP forecasts are adjusted such that the total of all regional forecasts meets the exogenous forecast of economic development (GDP) of the European Union as a whole by the *European Developments* submodel (see Section 2.4.1).

As it was explained in Section 2.3, the way of specifying and estimating the regional production functions will be further developed for the extended SASI model. Major changes will include:

- Instead of the original three, six economic sectors (the same ones as in the CGEurope model) will be considered: (i) Manufactured products, (ii) Market services, (iii) Agriculture, forestry and fishery products, (iv) Fuel and power products, (v) Building and construction and (vi) market services.
- It will be examined whether it is feasible of forecasting *rates* of change of regional economic development rather than the *levels* of regional production.

2.4.4 Regional Employment

Regional employment by industrial sector is derived from regional GDP by industrial sector and regional labour productivity.

Regional labour productivity was forecast in the original SASI model exogenously based on exogenous forecasts of labour productivity in each country:

$$p_{ir}(t) = p_{ir}(t-1) \frac{p_{ir'}(t)}{p_{ir'}(t-1)} \quad \text{with } r \in \mathbf{R}_{r'} \quad (2.9)$$

where $p_{ir}(t)$ is labour productivity, i.e. annual GDP per worker, of industrial sector i in region r in year t , $p_{ir'}(t)$ is average labour productivity in sector i in year t in country or group of regions $\mathbf{R}_{r'}$ to which region r belongs. The rationale behind this specification was the assumption that labour productivity by economic sector in a region is predominantly determined by historical conditions in the region, i.e. by its composition of industries and products, technologies and education and skill of labour and that it grows by an average sector-specific growth rate.

In the extended SASI model, the feasibility of forecasting regional sectoral labour productivity endogenously as a function of accessibility and other variables instead of using exogenous productivity forecasts will be explored. Because it can be assumed that labour productivity is positively affected by growth in accessibility, one possible specification might be

$$p_{ir}(t) = p_{ir}(t-1) \frac{p_{ir'}(t)}{p_{ir'}(t-1)} + \varepsilon_i \frac{A_r(t)}{A_r(t-1)} \quad \text{with } r \in \mathbf{R}_{r'} \quad (2.10)$$

where $A_r(t)$ is accessibility of region r in year t (aggregated across modes as above), and ε_i is a linear elasticity indicating how much the growth in labour productivity is accelerated by a growth in accessibility. As indicated above, absolute rather than relative accessibility is preferable here. Other specifications incorporating further variables may be explored.

Regional employment by industrial sector is then

$$E_{ir}(t) = Q_{ir}(t) / p_{ir}(t) \quad (2.11)$$

where $E_{ir}(t)$ is employment in industrial sector i in region r in year t , $Q_{ir}(t)$ is the GDP of industrial sector i in region r in year t and $p_{ir}(t)$ is the annual GDP per worker of industrial sector i in region r in year t .

2.4.5 Regional Population

The *Regional Population* submodel forecasts regional population by five-year age groups and sex through natural change (fertility, mortality) and migration. Population forecasts are needed to represent the demand side of regional labour markets.

Fertility and mortality

Changes of population due to births and deaths are modelled by a cohort-survival model subject to exogenous forecasts of regional fertility and mortality rates. To reduce data requirements, a simplified version of the cohort-survival population projection model with five-year age groups is applied. The method starts by calculating survivors for each age group and sex:

$$P'_{asr}(t) = P_{asr}(t-1) [1 - d_{asr}(t-1, t)] \quad \text{with } r \in \mathbf{R}_{r'} \quad (2.12)$$

where $P'_{asr}(t)$ are surviving persons of age group a and sex s in region r in year t , $P_{asr}(t-1)$ is population of age group a and sex s in year $t-1$ and $d_{asr}(t-1, t)$ is the average annual death rate of age group a and sex s between years $t-1$ and t in country or group of regions $\mathbf{R}_{r'}$ to which region r belongs.

Next it is calculated how many persons change from one age group to the next through ageing employing a smoothing algorithm:

$$g_{asr}(t-1, t) = 0.12 P'_{asr}(t) + 0.08 P'_{a+1sr}(t) \quad \text{for } a=1, 19 \quad (2.13)$$

where $g_{asr}(t-1, t)$ is the number of persons of sex s changing from age group a to age group $a+1$ in region r . Surviving persons in year t are then

$$P_{asr}(t) = P'_{asr}(t) + g_{a-1sr}(t-1, t) - g_{asr}(t-1, t) \quad \text{for } a=2, 19 \quad (2.14)$$

with special cases

$$P_{20sr}(t) = P'_{20sr}(t) + g_{19sr}(t-1, t) \quad (2.15)$$

$$P_{1sr}(t) = P'_{1sr}(t) + B_{sr}(t-1, t) - g_{1sr}(t-1, t) \quad (2.16)$$

where $B_{sr}(t-1, t)$ are births of sex s in region r between years $t-1$ and t :

$$B_{sr}(t-1, t) = \sum_{a=4}^{10} 0.5 [P'_{a2r}(t) + P_{a2r}(t)] b_{asr'}(t-1, t) [1 - d_{0sr'}(t-1, t)] \quad \text{with } r \in \mathbf{R}_{r'} \quad (2.17)$$

where $b_{asr'}(t-1, t)$ are average number of births of sex s by women of child-bearing five-year age groups a , $a = 4, 10$ (15 to 49 years of age) in country or group of regions $\mathbf{R}_{r'}$ to which region r belongs between years $t-1$ and t , and $d_{0sr'}(t-1, t)$ is the death rate during the first year of life of infants of sex s in country or group of regions $\mathbf{R}_{r'}$ to which region r belongs. The exogenous forecasts of death and birth rates in the above equations may be national rates or rates for specific groups of comparable regions.

Migration

In the original SASI model, both migration within the European Union and immigration from non-EU countries was modelled as annual regional net migration. For the extended SASI model, it is planned to forecast migration *flows* instead of *net* migration. A possible form of the migration model could be a singly constrained spatial interaction model

$$m_{rs}(t, t+1) = B_s L_r^{\alpha_r}(t) E_s^{\gamma_s}(t) \exp[-\beta c_{rs}(t)] \quad (2.19)$$

where $L_r(t)$ is labour in origin region r at time t , $E_s(t)$ is employment, or job opportunities, in destination region s at time t , and $c_{rs}(t)$ is generalised travel cost between regions r and s . The α_r and γ_s are push and pull factors, respectively:

$$\begin{aligned} \alpha_r &= f[u_r(t-1), v_r(t)] \\ \gamma_s &= f[u_s(t-1), v_s(t)] \end{aligned} \quad (2.20)$$

They are functions of unemployment $u_r(t-1)$ and quality of life $v_r(t)$ in the origin region and unemployment $u_s(t-1)$ and quality of life $v_s(t)$ in the destination region. The quality-of-life indicator $v_r(t)$ is also used as one of the endowment factors in the regional production function for service industries (see Section 2.4.3). Because at the time of execution regional unemployment in year t is not yet known, unemployment in the previous year $t-1$ is used.

The balancing factors B_s in Equation 2.19 guarantee that the total number of immigrants from non-EU countries to EU countries does not exceed the maximum number of immigrants set by the immigration policies of EU countries and forecast by the *European Developments* sub-model.

Educational attainment

Regional educational attainment, i.e. the proportion of residents with higher education in region r , is forecast exogenously assuming that it grows as in the country or group of regions to which region r belongs:

$$h_r(t) = h_r(t-1) h_{r'}(t) / h_{r'}(t-1) \quad \text{with } r \in \mathbf{R}_{r'} \quad (2.21)$$

where $h_r(t)$ is the proportion of residents with higher education in region r in year t , and $h_{r'}(t)$ is the average proportion of residents with higher education in country or group of regions $\mathbf{R}_{r'}$ to which region r belongs.

2.4.6 Regional Labour Force

Regional labour force is derived from regional population and regional labour force participation.

Regional labour force participation by sex is partly forecast exogenously and partly affected endogenously by changes in job availability or unemployment. It is assumed that labour force participation in a region is predominantly determined by historical conditions in the region, i.e. by cultural and religious traditions and education and that it grows by an average country-specific growth rate. However, it is also assumed that it is positively affected by availability of jobs (or negatively by unemployment):

$$\ell_{sr}(t) = \ell_{sr}(t-1) \ell_{sr'}(t) / \ell_{sr'}(t-1) - \varphi_s u_r(t-1) \quad \text{with } r \in \mathbf{R}_{r'} \quad (2.22)$$

where $\ell_{sr}(t)$ is labour force participation, i.e. the proportion of economically active persons of sex s of regional population of sex s 15 years of age and older, in region r in year t , $\ell_{sr'}(t)$ is average labour participation of sex s in year t in country or group of regions $\mathbf{R}_{r'}$ to which region r belongs, $u_r(t-1)$ is unemployment in region r in the previous year $t-1$ (see below), and φ_s is a linear elasticity indicating how much the growth in labour productivity is accelerated or slowed down by regional unemployment. Because at the time of execution of the *Regional Labour Force* submodel regional unemployment in year t is not yet known, unemployment in the previous year $t-1$ is used. Regional labour force by sex s in region r , $L_{sr}(t)$, is then

$$L_{sr}(t) = P_{sr}(t) \ell_{sr}(t) \quad (2.23)$$

where $P_{sr}(t)$ is population of sex s 15 years of age and older in region r at time t and $\ell_{sr}(t)$ is the labour force participation rate of sex s in region r in year t .

Regional labour force is disaggregated by skill in proportion to educational attainment in the region calculated in the *Population* submodel (see Section 2.4.5):

$$L_{sr1}(t) = h_r(t) L_{sr}(t) \quad (2.24)$$

with $L_{sr1}(t)$ being skilled labour and the remainder unskilled labour:

$$L_{sr2}(t) = L_{sr}(t) - L_{sr1}(t) \quad (2.25)$$

2.4.7 Socio-Economic Indicators

Total GDP and employment represent only the supply side of regional socio-economic development. To derive policy-relevant indicators, they have to be related to the demand side, i.e. to population and labour force. This is done by calculating total regional GDP per capita and regional unemployment. Since accessibility, besides being a factor determining regional production (see Section 2.4.3), is also an indicator of regional locational advantage and quality of life, accessibility indicators are considered a policy-relevant output of the model.

Accessibility, GDP per capita and unemployment are therefore the main socio-economic and spatial indicators produced by the SASI model. In addition, equity or *cohesion indicators* describing the distribution of accessibility, GDP per capita and unemployment across regions are calculated.

Accessibility

Regional accessibility indicators are calculated in the *Regional Accessibility* submodel (see Section 2.4.2)

GDP per capita

Despite its well-known theoretical and methodological drawbacks GDP per capita continues to be the most commonly used indicator of regional economic development. With certain qualifications, e.g. for regions with a large amount of commuting across their boundaries, GDP per capita allows to draw conclusions on regional income. Total regional GDP per capita is calculated as the sum of GDP by industrial sector divided by population:

$$q_r(t) = \sum_i Q_{ir}(t) / P_r(t) \quad (2.29)$$

Unemployment

Regional unemployment, too, presents measurement problems because there exist large differences in the definition of unemployment in European countries. Therefore the unemployment rate calculated in the SASI model only serves to compare different scenarios and is not comparable to the standardised unemployment rates calculated by Eurostat. Nevertheless unemployment remains the most widely used social indicator and is closely related to policy goals of the European Union.

To take account of interregional commuting, the resulting numbers of unemployed persons are reduced by outcommuters and increased by incommuters derived from a doubly constrained spatial-interaction work trip model. The work trip model is based on NUTS-3 shares of NUTS-2 population (see Fürst et al., 1999, Section 4.7). Before executing the work trip model, regional labour force is adjusted such that total labour force of all regions equals total employment of all regions.

The regional unemployment rate $u_r(t)$ in year t is

$$u_r(t) = \frac{L_r(t) - \sum_s T_{rs}(t) + \sum_s T_{sr}(t) - E_r(t)}{L_r(t)} \quad (2.30)$$

where $L_r(t)$ is total labour in region r in year t , $E_r(t)$ is total employment in region r in year t and $T_{rs}(t)$ are commuters from region r to region s in year t calculated from an attraction-constrained spatial-interaction work trip model:

$$T_{rs}(t) = \frac{L_r(t) \exp[-\beta c_{rs}(t)]}{\sum_r L_r(t) \exp[-\beta c_{rs}(t)]} E_s(t) \quad \text{with } \sum_s T_{rs}(t) \leq L_r(t) \quad (2.31)$$

where $c_{rs}(t)$ is travel time and/or cost between regions r and s in year t and the additional constraint ensures that there are no more workers than labour force in a region.

2.4.8 Cohesion Indicators

From the policy-relevant indicators so derived, equity or cohesion indicators describing their distribution across regions are calculated. Cohesion indicators are macroanalytical indicators combining the indicators of individual regions into one measure of their spatial concentration. Changes in the cohesion indicators predicted by the model for future transport infrastructure investments reveal whether these policies are likely to reduce or increase existing disparities in those indicators between the regions.

- *The coefficient of variation.* The coefficient of variation is the standard deviation of region indicator values expressed in percent of their European average:

$$V = \frac{\sqrt{\sum_r (\bar{X} - X_r)^2 / n}}{\bar{X}} 100 \quad (2.32)$$

The coefficient of variation informs about the degree of homogeneity or polarisation of a spatial distribution. A coefficient of variation of zero indicates that all areas have the same indicator values. The different size of regions can be accounted for by treating each area as a collection of individuals having the same indicator value. The coefficient of variation can be used to compare two scenarios with respect to cohesion or equity or two points in time of one scenario with respect to whether convergence or divergence occurs.

- *The Gini coefficient.* The Lorenz curve compares a rank-ordered cumulative distribution of indicator values of areas with a distribution in which all areas have the same indicator value. This is done graphically by sorting areas by increasing indicator value and drawing their cumulative distribution against a cumulative equal distribution (an upward sloping straight line). The surface between the two cumulative distributions indicates the degree of polarisation of the distribution of indicator values. The Gini coefficient calculates the ratio between that area of that surface and the area of the triangle under the upward sloping line of the equal distribution. The equation for the Gini coefficient is

$$G = 1 + 1/n - 2/(n^2 \bar{X}) \sum_r r X_r \quad (2.33)$$

where X_r are indicator values of regions r sorted in increasing order, $\bar{X}$ the average indicator value of all regions, and n the number of regions. A Gini coefficient of zero indicates that the distribution is equal-valued, i.e. that all areas have the same indicator value. A Gini coefficient close to one indicates that the distribution of indicator values is highly polarised, i.e. few areas have very high indicator values and all other areas very low values. The different size of areas can be accounted for by treating each area as a collection of individuals having the same indicator value.

A deficiency of both indicators is that they measure *relative* differences between spatial distributions. In a study for DG Regio (Wegener et al., 2001) it was found that the results of cohesion analyses heavily depend on the type of accessibility indicators applied. In the case of potential accessibility, the largest absolute gains in accessibility are concentrated in the most central and already most accessible areas, i.e. the existing disparities in accessibility increase. In relative terms, peripheral areas benefit more from better access to large central agglomerations. The conclusion was that relative cohesion indicators tend to suggest a tendency of convergence where in fact divergence may have occurred.

For the extended SASI model therefore the cohesion indicators to be used for assessing the impacts of transport policies will be expanded and critically assessed with respect to their possible implicit bias towards convergence and divergence.

2.5 Other Model Output

As indicated, the main output of the SASI model are accessibility, GDP per capita and unemployment for each regions for year of the simulation. However, a great number of other regional indicators are generated during the simulation. These indicators can be examined during the simulation or analysed and compared after several simulation runs.

During the simulation the user may monitor change processes in the model by observing time-series diagrams, choropleth maps or 3D representations of variables of interest on the computer display. The user may interactively change the selection of variables to be displayed during processing. The following options can be selected:

Population indicators

- Population (1981=100)
- Percent population 0-5 years
- Percent population 6-14 years
- Percent population 15-29 years
- Percent population 30-59 years
- Percent population 60+ years
- Labour force (1981=100)
- Labour force participation rate (%)
- Percent lower education
- Percent medium education
- Percent higher education
- Net migration per year (%)
- Net commuting (% of labour force)

Economic indicators

- GDP (1981=100)

- Percent non-service GDP
- Percent service GDP
- GDP per capita (in 1,000 Euro of 1998)
- GDP per capita (EU15=100)
- GDP per worker (in 1,000 Euro of 1998)
- Employment (1981=100)
- Percent non-service employment
- Percent service employment
- Unemployment (%)
- Agricultural subsidies (% of GDP)
- European subsidies (% of GDP)
- National subsidies (% of GDP)

Attractiveness indicators

- Accessibility rail/road (logsum, million)
- Accessibility rail/road/air (logsum, million)
- Soil quality (yield of cereals in t/ha)
- Developable land (%)
- R&D investment (% of GDP)
- Quality of life (0-100)

The same selection of variables can be analysed and post-processed after the simulation. If several scenarios have been simulated, the user can compare the results using a special comparison software (see Wegener et al., 2000). The selection of indicators available for output can be extended depending on the requirements of the present study, in particular in response to the needs of the CBA.

In addition to the regional indicators, travel cost matrices containing travel costs between all 1,336 European regions (see Section 4) will be calculated for use in the CGEurope model (see Section 3). The travel cost matrices for CGEurope are required for the benchmark year 1997 and the future year 2021.

2.6 Data Requirements

The original SASI model forecast socio-economic development in 201 regions at the NUTS-2 level for the fifteen EU countries. These were the 'internal' regions of the model. The 27 regions in 23 countries of the rest of Europe were 'external' regions which were used as additional destinations when calculating accessibility indicators. The extended system of regions to be used in the present study is described in Section 4. It has 1,083 regions at the NUTS-3 level in the present 15 EU member states, 162 regions of comparable size in the 12 candidate countries in eastern Europe and 91 regions in the remaining 14 countries of Europe. The 1,083+162 regions in the extended EU will be the 'internal' region of the extended SASI model. The 91 regions in the rest of Europe will be used as 'external' regions, i.e. as additional destinations for accessibility calculations.

In the extended SASI model, road, rail and air networks were considered. These networks have been updated and extended to link the extended system of regions. The updated and extended networks to be used in the present study are described in Section 4.

The base year of the simulations of the SASI model was 1981 in order to demonstrate that the model is able to reproduce the main trends of spatial development in Europe over a significant time period of the past with satisfactory accuracy. The forecasting horizon of the model was 2016. The base year of the extended SASI model will continue to be 1981, and the forecasting horizon will be 2021. This will allow twenty years of backcasting and twenty years of forecasting. For comparability with the CGEurope model, output for the benchmark year of the CGEurope model, 1997, will be provided.

This section describes the data required for the extended SASI model from the Common Database to be used by both models. The specification of the Common data Base will be presented next. Two major groups of data are distinguished: data required for running the model (simulation data) and data needed for the calibration or validation of the model. In each of these categories, the data can be classified by spatial and temporal reference.

2.6.1 Simulation Data

Simulation data are the data required to perform a typical simulation run. They can be grouped into *base-year* data and *time-series* data.

Base-year data

Base-year data describe the state of the regions and the strategic road, rail and air networks in the base year. Base-year data are either regional or network data.

Regional base-year data are required to provide base values for the *Regional GDP* submodel and the *Regional Population* submodel as well as base values for exogenous forecasts of changes in regional educational attainment and regional labour force participation. All other regional base-year values such as GDP, employment or labour force are calculated by the model (even where regional base-year data for these variables are available). Network base-year data specify the road, rail and air networks used for accessibility calculations in the base year.

Regional data (1,083+162 regions, see Section 4)

- Regional GDP per capita by industrial sector in 1981
- Regional labour productivity (GDP per worker) by industrial sector in 1981
- Regional population by five-year age group and sex in 1981
- Regional educational attainment in 1981
- Regional labour force participation rate by sex in 1981
- Regional quality-of-life indicators in 1981

Network data

- Node and link data of strategic road network in 1981
- Node and link data of strategic rail network in 1981
- Node and link data of air network in 1981

In addition, for the allocation of regional population and GDP to raster cells in the *Regional Accessibility* submodel, the disaggregate raster distributions of population and GDP are required. For simplicity, population and GDP of the most recent available year are used for the distribution in all years.

Time-series data

Time-series data describe exogenous developments or policies defined to control or constrain the simulation. They are either collected or estimated from actual events for the time between the base year and the present or are assumptions or policies for the future. Time-series data must be defined for each simulation period, but in practice may be entered only for specific (not necessarily equidistant) years, with the simulation model interpolating between them. All GDP data are converted to Euro of 2001.

- European data (15+12 countries, see Section 4)

 - Total European GDP by industrial sector, 1981-2021

 - Total European immigration and outmigration, 1981-2021

- National data (15+12 countries, see Section 4)

 - National GDP per worker by industrial sector, 1981-2021

 - National fertility rates by five-year age group and sex, 1981-2021

 - National mortality rates by five-year age group and sex, 1981-2021

 - National immigration limits, 1981-2021

 - National educational attainment, 1981-2021

 - National labour force participation by sex, 1981-2021

- National data (14 non-EU countries, Section 4)

 - National population, 1981-2021

 - National GDP, 1981-2021

- Regional data (1,083+162 regions, see Section 4)

 - Regional endowment factors, 1981-2021

 - Regional transfers, 1981-2021

- Network data

 - Changes of node and link data of strategic road network, 1981-2021

 - Changes of node and link data of strategic rail network, 1981-2021

 - Changes of node and link data of air network, 1981-2021

2.6.2 Calibration/Validation Data

The regional production function in the *Regional GDP* submodel and the migration function in the *Regional Population* submodel are the only model functions *calibrated* using statistical estimation techniques. All other model functions are *validated* by comparing the output of the whole model with observed values for the period between the base year and the present.

Calibration data

Calibration data are data needed for calibrating the regional production functions in the *Regional GDP* submodel, labour productivity in the *Regional Employment* submodel and the migration function in the *Regional Population* submodel. The three calibration years 1981, 1986, 1991 and 1996 are suggested to gain insights into changes in parameter values over time; however, the calibration is also possible with less calibration years. The calibration data of 1981 are partly identical with the simulation data for the same year.

- Regional data (1,083+162 regions)

 - Regional GDP per capita by industrial sector in 1981, 1986, 1991, 1996

- Regional labour productivity by industrial sector in 1981, 1986, 1991, 1996
- Regional wage levels and/or production costs by sector in 1981, 1986, 1991, 1996
- Regional endowment factors in 1981, 1986, 1991, 1996
- Regional labour force in 1981, 1986, 1991, 1996
- Regional transfers in 1981, 1986, 1991, 1996
- Regional migration flows in 1981, 1986, 1991, 1996
- Regional unemployment rates in 1981, 1986, 1991, 1996

Network data

- Node and link data of strategic road network in 1981, 1986, 1991, 1996
- Node and link data of strategic rail network in 1981, 1986, 1991, 1996
- Node and link data of air network in 1981, 1986, 1991, 1996

Validation data

Validation data are reference data with which the model results in the period between the base year and the present are compared to assess the validity of the model. Validation is preferable over calibration where processes simulated in the model are unobservable or unobserved because of lack of data. Validation can be used to experimentally adjust model parameters that cannot be calibrated until the model results match available aggregate data. The validation years suggested below are indicative; the validation can be performed with less observations. Also the disaggregations indicated in brackets are optional.

Regional data (1,083+162 regions)

- Regional population (by age and sex) in 1981, 1986, 1991, 1996, 2001
- Regional GDP (by industrial sector) in 1981, 1986, 1991, 1996, 2001
- Regional labour force (by sex) in 1981, 1986, 1991, 1996, 2001
- Regional employment (by industrial sector) in 1981, 1986, 1991, 1996, 2001
- Regional unemployment rate in 1981, 1986, 1991, 1996, 2001

3 The New CGEurope Model

3.1 Modelling the Spatial Dimension of Transport Benefits

Although extensive research is already under way for assessing the infrastructure needs as well as costs and benefits of individual projects, very little is still known about the spatial distribution of the benefits. Traditional approaches to cost benefit and regional impact analysis are not really capable of taking account of the complex mechanisms by which transport cost changes affect the spatial allocation. This holds true already in a static framework, not to speak about the even more complex channels through which the transport system aspects economic dynamics. The critical issue is to assign the benefits from using the transport links to regions. Assigning costs and benefits from construction and maintenance to regions is less of a problem, and traditional techniques like multiplier analysis are acceptable. Assessing the benefits from newly installed capacities and answering the question where they accrue, however, is much more difficult. Four types of methods are used in practice:

The *first* is to assign benefits as measured by direct cost reductions or consumer surpluses gained on the links under study to the place of investment itself. This method is applied in the official German manual for transport infrastructure evaluation (BMV, 1993), for example. Its shortcomings are so obvious that a further discussion is not worth the effort.

The *second* method is to measure benefits by estimating rates of return on infrastructure investments in a production function approach, using cross section, time series, or panel data (for a survey see Pfähler et al., 1996). Intricate econometric problems have to be solved for this type of analysis, which are thoroughly discussed in an extensive literature. As far as the regional distribution of effects is concerned, however, the shortcomings of this approach are similar to those of the first one. While accessibility changes may affect many regions possibly in a different way depending on the pattern of inter-regional flows, all output effects are exclusively attributed to the region, where the investment is done.

The *third* method is to establish an inter-regional demand-driven input-output model with trade coefficients depending on transportation costs (see e.g. Leitham, Downing, Martino, and Fiorello 1999). Though this seems attractive because a lot of sectoral detail can be taken account of, it gives a theoretically unconvincing picture of the effects of changing transport costs. It is restricted to backward linkage effects. In this type of approach, it is difficult to simulate the cost effects and price effects stemming from a reduction in transport costs. To extend the picture to forward linkages generated by increased product diversity brought about by integrating the local markets is even more difficult.

The *fourth* method is to measure the impact of transport cost reductions by accessibility indicators telling how a region's generalised cost of reaching its markets and travelling to a hypothetical set of destinations is affected by the cost reductions (see Vickerman et al., 1999 for an empirical example and Rietveld and Bruinsma, 1998a for a survey). In a second step, accessibility changes are then related to regional economic indicators like GDP per capita or real growth of GDP, using cross-section regression techniques. This is the basic idea of the SASI approach and is discussed in Section 2 in more detail.

The *fifth* more recent technique is to set up a multi-regional computable general equilibrium, in which transport costs explicitly appear as firms' expenditures for transport and other kinds of business travel and as households' costs of private passenger travel (for examples see Venables and Gasior, 1998, Bröcker, 1998a). This is what we are doing with the CGEurope model (see Bröcker, 1998a, 1999, 2000, 2001). CGEurope is a multiregional, and in its extended version developed for the present study multi-sectoral, computable general equilibrium model, incorporating innovative features from recent developments in the literature like product diversity and monopolistic competition, explicit modelling of out-of-pocket as well as time costs of business transport as well as private passenger transport.

The way, how transport cost changes are modelled in this framework is obvious. After having calibrated the model such that the data of a benchmark year are reproduced, transport costs or travel times are changed exogenously and the new equilibrium system is solved. The main indicator for the regional consequences one is looking at is the welfare change of regional households, as measured by the households' utility functions. Though an ordinal utility index as it stands has no operational meaning, it can be transformed to the so-called Hicks measures of variation. They measure the welfare change in monetary terms.

CGEurope is confined to the regional welfare effects resulting from the use of the transport infrastructure. Effects from the construction phase, from financing and maintenance are not considered. We also do not include local traffic including commuting, even if it is commuting over longer distances crossing the borders of the regions in our system.

3.2 Extending CGEurope

Compared to the previous version of CGEurope (for a description see Bröcker, 1998a) the new version for the present study is extended in the following respects:

- The previous version had only two sectors (tradable and non-tradable), while the new one differentiates between six sectors, including one sector producing the transport service using factors and intermediate inputs (for a detailed description see Section 3.5).
- The previous version took only transport costs in interregional trade into account, while the new one also includes costs of private passenger travel.
- The new version of CGEurope models the use of resources for transport in a more sophisticated way than the previous one by including explicitly an activity producing the transport service.

Finally, the transport network from which the cost measurement is derived is much more refined, based on the networks developed within SASI, SCENES and ETIS.

For the present study we have developed the new model version in a way that allows for calibration with existing data. A detailed description of the new version of CGEurope is given in section 3.3. In particular, the following tasks have been successfully carried out in Work Package 2.1:

Sectoral and regional coverage

Definition of *new sectors and regions*. The previous version of CGEurope covered 805 regions for one tradable and one non-tradable manufacturing sector only and was based on Eurostat's Regional Accounts. The extended version developed over the last seven months covers six activities (including services) and a wider range of regions – 49 countries and country groups and 1,341 regions (for a detailed description see Section 4). Taking data availability into consideration, however, it will not be possible to have results with full sectoral detail and full regional detail at the same time. It is therefore necessary to run two different versions of the model, one with aggregated regions, and another with full regional detail but aggregated sectors.

Model structure

Setting up the *system of equations describing the multi-sectoral system* of the new model. The main problem, which had to be solved, was to design the model in a way allowing calibration with limited information on a sub-national regional scale (see Section 3.4).

Travel demand

Developing a new approach to model passenger *travel behaviour in a microeconomically consistent framework* that can consistently be integrated into the general equilibrium context. In particular, we had to include monetary travel costs as well as time costs into this framework, because time costs are an important determinant of travel behaviour and the change of time costs is an essential element of households' welfare. We succeeded in fulfilling three requirements, namely (1) to derive behaviour and welfare measures from one single theoretical formulation, (2) to specify travel preferences in a way that observed dispersed travel behaviour can realistically be reproduced, and (3) to make things sufficiently simple such that the parameters can be calibrated with minimum data requirements. We need data for expenditure shares for interregional travel (excluding commuting and other kinds of local travel) and for interregional travel flows in quantity terms. The latter may also be rough estimates based on gravity-like hypotheses. For a rough technical explanation of the theory of household behaviour applied in the new model version, see the two following sections.

Calibration

Developing a *calibration procedure working with a limited database* without a full multiregional social accounting matrix. In this respect, our approach deviates from available work in the fields of computable general equilibrium modelling. Usually one has original or derived full information about monetary flows between each agent (firm or household) in each region for the benchmark year. This covers trade by sector between firms, trade between firms and households, factor expenditures flowing from firms to households and interregional capital flows. As described in Section 3.5, most probably it will be impossible to obtain a full database at a sub-national (NUTS-2) level. Hence, we have developed a different approach, effectively combining information on the distribution of sectoral output by region with national and international information on national accounts and international trade (see the following section). We assume identical preferences and technologies for different regions within one

nation, such that national information is sufficient for calibrating technology parameters (for a full technical description see Bröcker, 1995, 1998a and 1998b).

Interregional trade on the sub-national level is not observed either, but derived from the calibrated equilibrium solution. The essential hypothesis in this context is that customers of traded goods substitute between varieties stemming from different regions, taking prices and interregional transaction costs into account. These transaction costs also include international trade impediments (cross border effects), which are indirectly quantified by adjusting estimated trade flows to the international totals available from international trade statistics. Even though these calibration techniques have already been used in the former CGEurope, applying them in the extended multi-sectoral framework is much more complicated, and we had to set up the nonlinear system of calibration equations needed to solve this problem. The solution algorithms for this system envisaged still wait for their test in the large real world application.

3.3 Non-Formal Description of the New CGEurope

CGEurope is a multiregional model for a closed system of regions, treating separately each region and linking them through endogenous trade. The world is subdivided into a large number of regions.¹ Each region shelters a set of households owning a bundle of immobile production factors used by regional firms for producing goods and services. The new version of CGEurope distinguishes six different sectors, five of which are tradable and one non-tradable (local) good (see Section 3.5). Beyond factor services, firms also use local goods and tradables as inputs. The firms in a region buy local goods from each other, while tradables are bought everywhere in the world, including the own region. Produced tradables are sold everywhere in the world, including the own region. Free entry drives profits to zero; hence, the firms' receipts for sold local goods and tradables equal their expenditures for factor services, intermediate local and tradable goods and transport.

Regional final demand, including investment and public sector demand, is modelled as expenditure of utility maximising regional households, who spend their total disposable income in the respective period. Disposable income stems from returns on regional production factors, which, by assumption, are exclusively owned by regional households, and a net transfer payment from the rest of the world. This transfer income can be positive or negative, depending on whether the region has a trade deficit or surplus. Transfers are held constant in our simulations. Introducing fixed interregional income transfers is a simplified way to get rid of a detailed modelling of interregional factor income flows, and of all kinds of interregional flows of private and public funds. Households expend their income for local and tradable goods as well as for travel. The vector of travel demand is differentiated by purpose of travel and destination. Households gain utility from a set of activities connected with travel (like tourism) and suffer from disutility for spending travel time.

¹ The new version of CGEurope covers 1,341 regions, of which 1,336 cover Europe, including the Asian parts of Russia and Turkey. In addition, there are 5 non-European regions covering the rest of the world, namely North America, Latin America, Africa, Middle East, and Asia plus Australia and New Zealand.

How many primary factors will be introduced is still open; it depends on the detail of information available in the national accounts. One version will be with just one homogeneous factor² Differentiation between factors depends on data availability.

The factor supply is always fully employed due to the assumption of perfect price flexibility, which implies the assumption that the rate of unemployment remains unaffected by the exogenous influences under study. Analysing effects on unemployment requires a deeper study of the structure of labour markets, which is not part of this project. We assume complete immobility of factors, which means that interregional factor movements as a reaction to changing transport costs is not included. The other extreme assumption would be perfect factor mobility, but this is not realistic. Immobility is taken as a first approximation for short-term effects. The best choice would be mobility, but an imperfect one. There are ways of introducing such an assumption, but theoretically consistent approaches require forward-looking dynamics, which are too complicated to be introduced into our model in the present stage of its development.

Firms representing production sectors are of two kinds, producers of local goods and producers of tradables. Each local good is a homogeneous good, though one equivalently may regard it as a given set of goods, such that the good's price is to be interpreted as the price of a composite local good. The market for tradables, however, is modelled in a fundamentally different way. Tradables consist of a large number of close but imperfect substitutes. The set of goods is not fixed exogenously, but it is determined in the equilibrium solution and varies with changing exogenous variables. Different goods stem from producers in different regions. Therefore, relative prices of tradables do play a role. Changes of exogenous variables make these relative prices change and induce substitution effects.

Households act as price taking utility maximisers. They have a nested CES utility function representing substitution between goods and travel activities, between goods from different sectors, between different kinds of travel activities, between destinations for each kind of travel and between varieties for each kind of goods. In the disutility version for modelling the burden of travel time, a travel time disutility is subtracted from the households' utility function in an additive separable format.

Firms maximise profits. Local goods producers take prices for inputs as well as for local goods sold to households and other firms as given. The production functions are linear-homogenous nested-CES functions. The lowest CES nest makes a composite out of the bundle of tradables. For the sake of simplicity, it is assumed to be identical for all users and to be the same as the respective CES nest in the households' utility function. Due to linear homogeneity, the price of local good equals its unit cost obtained from cost minimisation under given input prices.

Tradable goods producers take only prices for inputs as given. They produce a raw output by a technology designed in the same way as for local goods producers. Instead of directly selling their output, however, they transform the homogeneous raw output into a final differenti-

² The reader should be aware that we equivalently could regard factor supply as a vector of arbitrary length representing labor with different qualifications, capital, land etc. The factor price has then to be interpreted as a price for a composite factor service. The important assumption is that with just one composite factor in the model we exclude sectoral variation of factor intensities, such that relative factor prices are not going to play a role in the solution

ated output. The respective technology is increasing returns, with a decreasing ratio of average to marginal input. Firms are free to compete in the market for a tradable good, which already exists, or to sell a new one not yet in the market. The latter turns out to be always the better choice. Hence, only one firm monopolistically supplies each good, which is aware of the finite price elasticity of demand for the good. The firm therefore sets the price according to the rules of monopolistic mark-up pricing. This choice, of course, is only made if the firm at least breaks even with this strategy. If it comes out with a positive profit, however, new firms are attracted opening new markets, such that demand for each single good declines until profits are driven back to zero.

This is the well-known mechanism of Chamberlinian monopolistic competition determining the number of goods in the market as well as the quantity of each single good (see Krugman, 1993, Fujita et al., 1999, Bröcker, 1998a). Due to free entry, the price of a tradable good just equals its average unit cost. It turns out that under the assumption of a constant price elasticity of demand for each variety of goods, which is valid in our framework, output per variety is also constant, such that output variations come in the form of variations in the number of varieties, and real output is the endogenous measure of variety.

Certainly, assuming local markets to be perfectly competitive lacks empirical plausibility. Local goods producers may in fact exert some monopoly power, local goods might be diversified, just like tradables, et cetera. The reason why this assumption is nevertheless preferred is that this is the simplest way to get rid of the local sectors, which only play a secondary role in an analysis focusing on interregional trade. Another choice without major technical problems would be to assume monopolistic competition for the local sectors as well. This, however, is not recommended, because it introduces a size-of-region effect. Large regions in our system (like the Asian part of Russia, for example) would support a high diversity of local goods, generating an unrealistic low prices of composite local goods, given the factor price(s) and technology in the region.

Three features give the CGEurope model its spatial dimension:

- the distinction of goods, factors, firms and households by location,
- the explicit incorporation of transport cost for goods (and services, regarded as a special kind of goods), depending on geography as well as national segmentation of markets, and
- the explicit incorporation of private passenger travel, with time costs and out-of-pocket costs depending on geography as well as national segmentation of space.

Summarising the basic philosophy of our approach, it obviously strongly relies on neo-classical ideas, even though it departs from the traditional computable general equilibrium approach by allowing for imperfect markets. In other respects, however, the strictness of neo-classical assumptions is retained: firms and households act perfectly rationally, prices are flexible, and markets are cleared, including labour markets. Though these assumptions are often criticised for contrasting with reality, there is no better choice. Even if households don't maximise utility subject to a budget constrained, it is not questioned that they react on prices and that the budget constraint must eventually hold. Neo-classical demand theory is just an easy way to represent these reactions consistently in a formal way. Similar comments apply to modelling reactions of firms.

The issue is not whether the model is close to reality; no model will ever be so. The issue is which is the best way to represent fundamental mechanisms detected by theory in a quantita-

tive approach. In this context, marginal returns of making a model more complicated have to be traded off against marginal costs. More realistic models like large-scale econometric or input-output models with many sectors might offer a more realistic description, but are much more expensive and offer less possibilities for studying the interaction between prices and quantities in a theoretically consistent framework.

3.4 The Mathematical Structure of the New CGEurope Model

3.4.1 The Equations of the Equilibrium

We start with notation. Subscripts $r = 1, \dots, R$ denote regions, superscripts i or $j = 1, \dots, I$ denote sectors, and superscripts $k = 1, \dots, K$ denote factors. By convention, sector 1 is the sector producing the transport service. Finally, superscripts $\ell = 1, \dots, L$ denote travel activities.

- A_r $(I \times I)$ -matrix of intermediate input-coefficients with typical entry a_r^{ij} denoting the input of goods from sector i per unit of sector j 's output. If i denotes a tradable, the respective input is meant to be the CES-composite made of all the varieties bought in region r as well as in all other regions. This composite is the same for firms using it as an input as for households consuming it, as already mentioned
- B_r $(K \times I)$ -matrix of primary input-coefficients with typical entry b_r^{kj} denoting the input of factor k per unit of sector j 's output
- X_r $(I \times 1)$ -vector of regional outputs with typical entry X_r^i
- p_r the corresponding price vector with typical entry p_r^i
- D_r $(I \times 1)$ -vector of regional demand for (composite) goods with typical entry D_r^i
- q_r the corresponding price vector with typical entry q_r^i
- F_r $(I \times I)$ -vector of regional final demand for (composite) goods
- $\tilde{F}_r$ $((I \times I) \times 1)$ -vector of regional final demand for (composite) goods, with first element (demand for transport service) deleted
- f_r $(K \times I)$ -vector of regional factor demand with typical entry f_r^k
- S_r $(K \times I)$ -vector of regional factor supply with typical entry S_r^k
- w_r the corresponding price vector with typical entry w_r^k
- t_{rs}^i tradables from sector i delivered from region r to region s
- τ_{rs}^i mark-up for transport costs; the costs for shipping one unit of i -goods from region r to region s is $p_{rs}^i (\tau_{rs}^i - 1)$, such that the price for this unit in region s equals $p_{rs}^i \tau_{rs}^i$
- σ_i elasticity of substitution between varieties produced by sector i
- $c_r^i(\cdot)$ unit cost function derived from cost minimisation subject to the representative firm's technology; the function's argument is the vector of input prices $(q_r' w_r')^3$, its value is the cost per unit of sector i 's output

³ Apostroph ' denotes transposition

- $\mathbf{y}_r$ $(L \times 1)$ -vector of households' travel activities with typical entry y_r^ℓ denoting the travel demand of kind ℓ by households living in region r ; a kind of activity is travel for a certain purpose to a certain destination like tourism travel to destination s , say
- $\boldsymbol{\theta}_r$ $(L \times 1)$ -vector of travel times, with typical entry θ_r^ℓ denoting the travel time required per unit of activity of kind ℓ by households living in region r
- $\boldsymbol{\mathcal{G}}_r$ $(L \times 1)$ -vector of travel services, with typical entry $\mathcal{G}_r^\ell$ denoting the amount of travel service required per unit of activity of kind ℓ by households living in region r
- N_r net income transfers from other regions to region r
- $\mathbf{d}_r(\cdot)$ household's behaviour function, assigning vectors of final goods demand, factor supply and passenger travel demand to net income transfer and to vectors of prices, transport costs and travel times

The general equilibrium of the multiregional world economy is summarised by the following system of equations, which we will explain step by step:

$$\mathbf{D}_r = \mathbf{A}_r \mathbf{X}_r + \mathbf{F}_r \quad (3.1)$$

$$\mathbf{f}_r = \mathbf{B}_r \mathbf{X}_r \quad (3.2)$$

$$a_r^{ij} = \frac{\partial c_r^j(\mathbf{q}_r, \mathbf{w}_r)}{\partial q_r^i} \quad (3.3)$$

$$b_r^{kj} = \frac{\partial c_r^j(\mathbf{q}_r, \mathbf{w}_r)}{\partial w_r^k} \quad (3.4)$$

$$p_r^i = c_r^i(\mathbf{q}_r, \mathbf{w}_r) \quad (3.5)$$

$$q_s^i = \begin{cases} \left(\sum_r \psi^i X_r^i (p_r^i \tau_{rs}^i)^{1-\sigma^i} \right)^{\frac{1}{1-\sigma^i}} & \text{if } i \text{ is tradable,} \\ p_s^i & \text{else,} \end{cases} \quad (3.6)$$

$$t_{rs}^i = \begin{cases} \psi^i X_r^i \left(\frac{q_s^i}{p_r^i \tau_{rs}^i} \right)^{\sigma^i} D_s^i & \text{if } i \text{ is tradable,} \\ \begin{cases} D_s^i & \text{for } r = s \\ 0 & \text{else} \end{cases} & \text{else,} \end{cases} \quad (3.7)$$

$$\begin{pmatrix} \tilde{\mathbf{F}}_r \\ \mathbf{y}_r \end{pmatrix} = \mathbf{d}_r(\mathbf{q}_r, \boldsymbol{\theta}_r, \boldsymbol{\mathcal{G}}_r, Y_r) \quad (3.8)$$

$$\mathbf{F}_r^1 = \mathbf{S}'_r \mathbf{w}_r + N_r \quad (3.9)$$

$$\mathbf{F}_r^1 = \boldsymbol{\theta}'_r \mathbf{y}_r + \frac{1}{q_r^1} \sum_i \sum_s (\tau_{sr}^i - 1) p_s^i t_{sr}^i \quad (3.10)$$

$$\mathbf{X}_r^i = \sum_s t_{rs}^i \quad (3.11)$$

$$S_r = f_r. \quad (3.12)$$

Equations 3.1 and 3.2 are the familiar input output equations of a Leontief system, giving the total (intermediate plus final) demand for goods and factor demand. As already noted, the same composition index for varieties is assumed for firms and households, such that their respective demand for a composite good can be merged into one aggregate regional demand for that good.

Equations 3.3 and 3.4, stating Shephard's Lemma, give the goods-input and factor-input coefficients. Note that coefficients are endogenous, in general, but a Leontief technology with fixed coefficients is allowed as a special case. Equation 3.5 states that the output price equals minimal unit costs.

Equations 3.6 and 3.7 need some explanation. As far as tradables are concerned, the equations are implied by the assumption that varieties are merged to a composite good by a symmetrical CES index, as well as by the fact that the number of brands offered by industry i in region r is X_r^i (times a constant factor hidden in the parameter ψ^i , which just fixes units of measurement). Strictly speaking the number of brands is an integer. But it is treated as a real variable here, which is justified if the number of brands is large. Due to symmetry all brands from region r have the same fob price p_r^i . Thus, their cif price in s equals $p_r^i \tau_{rs}^i$. Inserting this into the CES demand system yields 3.6 and 3.7 for the case of tradables⁴. The non-tradables case is obvious: demand is served by the region itself only, without interregional trade.

Equation 3.8 models household behaviour. Households in region r decide upon goods consumption $\tilde{F}_r$ and passenger travel y_r . The decision depends on goods prices q_r , on the monetary and time costs of passenger travel, w_r and θ_r , and on income Y_r , which is factor income plus net transfer received (Equation 3.9). Note that the consumption vector excludes consumption of transport service. Demand for transport service is a derived demand. It is given by Equation 3.10, saying that final demand for the transport service is transport service required by households plus transport service required by firms. The latter term is absent if goods transport uses the composite good instead of a special transport service, as explained in Footnote 4.

Finally, equations 3.11 and 3.12 are the equilibrium conditions requiring clearing of goods and factor markets.

3.4.2 Household Behaviour and Welfare Measurement

⁴ Here it is assumed that the transport costs are expended for transport services produced in a special sector located in the region of destination. In the current version of CGEurope the transport service is produced using the composite itself as input. We will test this specification as well to see whether there are any important differences. With this specification the formula for t_{rs}^i looks slightly different, namely

$$t_{rs}^i = \frac{X_r^i (p_r^i \tau_{rs}^i)^{-\sigma^i}}{\sum_r X_r^i p_r^i (p_r^i \tau_{rs}^i)^{-\sigma^i}} D_s^i q_s^i$$

D_s^i is demand without the part required for transport (see Bröcker, 1998a for a derivation).

Households are assumed to maximise utility subject to their respective budget constraint. The utility $u(\tilde{\mathbf{F}}_r, \mathbf{y}_r, \theta'_r)$ depends on goods consumption, passenger travel and travel time, with marginal utility of travel time assumed to be negative, of course. The utility will be assumed to have a nested CES form with respect to the two consumption vectors $\tilde{\mathbf{F}}_r$ and $\mathbf{y}_r$ and to be additive with respect to travel time (dis-)utility. The weight of the travel time term is obtained from the empirical literature quantifying willingness to pay for time savings in passenger travel. It may also be allowed to insert weights varying over travel purposes, if that is indicated by the available willingness to pay estimates. The household's budget constraint is

$$Y_r = \tilde{\mathbf{F}}_r' \tilde{\mathbf{q}}_r + q_r^1 \theta_r^1 \mathbf{y}_r \quad \text{with } \tilde{\mathbf{q}}_r = (q_r^2, \dots, q_r^i)'$$

The consistent micro-foundation of household behaviour allows for a welfare measurement fully in line with the ideas of cost-benefit analysis. Welfare effects of any exogenous change like a decrease in transport costs are measured by comparing utility levels by region before and after the change. As utility levels have no meaning in a metric sense (they measure only on an ordinal scale), utility changes are translated to equivalent monetary values by Hicks' concept of equivalent variation (EV). Let us call the situation before the change the benchmark, and the situation after the change the alternative. The EV of the respective change is defined as the amount of money one would have to add to the household's benchmark income (everything else held constant on benchmark levels) in order to make the household as well off as in the alternative. Note that EV is not the same as the income increase generated by the change. This would be so only if no variable influencing utility but income changed. Other variables like prices and travel times do change, however, as a consequence of e.g. transport infrastructure investments. Regional EVs can be reported as per capita amounts and as shares in benchmark regional GDP.

3.4.3 Transport Costs

Now we show formally how transport costs are introduced. We start with trade costs and then briefly mention the specification of travel costs for private passenger travel.

The term "transport cost" for interregional trade is used as a shortcut for any kind of trade related costs. Usually trade costs are assumed to depend on the quantity of goods traded. Some costs of interregional transfer, especially costs of information exchange and insurance costs, depend on the value rather than the quantity traded, however. Letting trading costs depend on the value of trade makes the model much simpler, and we therefore prefer this assumption.

We introduce two kinds of trade costs: costs related to geographic distance, and costs for overcoming impediments to international trade. If region r belongs to country m and region s to country n , then the mark up factor (omitting the index i) is

$$\tau_{rs} = \phi(g_{rs}) \delta_{mn}. \quad (3.13)$$

g_{rs} denotes generalised transport distance quantified by the SASI model. ϕ is the transport cost function with $\phi(g_{rs}) = 1$. A plausible assumption is that ϕ increases with increasing distance, but at a diminishing rate. An obvious specification would be

$$\tilde{\phi}(g_{rs}) = 1 + \zeta(g_{rs})^\omega$$

with parameters $\zeta > 0$ and $0 < \omega < 1$.

The problem with this specification, however, is the following. The parameters of the transport cost function will be estimated using observations on international trade. It turns out that the cost function appears in a gravity formula for interregional trade in the equilibrium solution. The gravity equation has the distance function $[\tilde{\phi}(g_{rs})]^{-\sigma}$, which has to be fitted to observed trade patterns. Unfortunately it is impossible to estimate the three parameters σ, ζ, ω appearing in this function, because the effects of two of them, σ and ζ , are not separable from one another. Technically speaking, the level sets of the likelihood function are close to degeneration in (σ, ζ) -space. The reason is easy to see. If $\tilde{\phi}$ is sufficiently close to one (in the order of 1.2, say), then

$$\tilde{\phi}(g_{rs}) \approx \exp[\zeta(g_{rs})^\omega]$$

and

$$[\tilde{\phi}(g_{rs})]^{-\sigma} \approx \exp[-\sigma\zeta(g_{rs})^\omega].$$

Here ζ and σ merge to a single parameter $\sigma\zeta$. Hence, we prefer the specification

$$\phi(g_{rs}) = \exp[\zeta(g_{rs})^\omega], \tag{3.14}$$

implying that the gravity distance function becomes exactly

$$[\phi(g_{rs})]^{-\sigma} = \exp[-\sigma\zeta(g_{rs})^\omega]$$

ω and the merged parameter $\sigma\zeta$ are now well estimable, and we need other sources of information to separate the estimates for ζ and σ . See Section 3.2 for more on this.

ϕ is not globally concave. Its second derivative is negative for small g_{rs} , but changes sign for a sufficiently large g_{rs} . For the specific parameters we are working with, however, ϕ remains concave even for the longest distance in our system, and ϕ and $\tilde{\phi}$ do not differ much.

$(\delta_{mn} - 1) \geq 0$ is a tariff equivalent of all costs stemming from the fact that a good has to be exported from country m to country n . These include tariffs, but also, and more important, all costs stemming from non-tariff barriers, like costs due to language differences, costs for bureaucratic impediments, time costs spent at border controls and so forth. $\delta_{mn} = 1$ for $m = n$, of course, but it is suggested to be strictly larger than unity for $m \neq n$, even if countries m and n are both members of the EU (for a recent survey on international trade barriers see Bröcker, 1990 and Helliwell, 1998).

For passenger transport we do not need mark-up factors, but transport costs per unit of activity. The out-of-pocket costs are likely close to linear or slightly concave with respect to distance. Hence, if y_r^ℓ means travel volume for a certain purpose to destination s , say, then

$$g_{rs}^\ell = \alpha \cdot g_{rs}^\epsilon$$

with parameters $\alpha > 0$ and $0 < \epsilon < 1$ varying over purposes, but not over destinations.

3.4.4 Calibration

Calibrating the model means to assign concrete numbers to each parameter and exogenous variable such that the equilibrium solution exactly reproduces the observed data or resembles them as closely as possible. Unfortunately, however, this cannot provide all required parameters. In particular, fixing elasticities of substitution has to rely largely on literature surveys. The parameters to be chosen are:

1. Position parameters (also called shift parameters) in the cost functions $c_r^i(\cdot)$. We have not sufficient information for specifying them on a regional level. Hence, we assume them to be identical for all regions in a country, except that Harrod-neutral regional productivity levels are allowed within each sector. The position parameters are calibrated such that the input values in the national aggregates and the regional output values as reported in the social accounting system are reproduced in the benchmark equilibrium solution.
2. Position parameters in the behaviour function d_r of households. These concern goods and travel. Regarding goods, the position parameters are derived from reproducing final demand in the national accounts, assuming identical goods preferences over regions. Regarding travel, the position parameters are chosen such that travel information (or estimates) in the SCENES data base are reproduced.
3. Mark-up factors for international trade impediments δ_{mn} , $m \neq n$, are calibrated such that observed international trade flows equal the corresponding aggregates of trade flows between the regions of the two respective countries.
4. Elasticities of substitution in $c_r^i(\cdot)$ and d_r are borrowed from the extensive literature (Jomini et al., 1991, Hansen and Heckman, 1996, Shiells et al., 1986).
5. Valuations of time needed to calibrate d_r are also taken from the literature (Gálvez and Jara-Díaz, 1998, Jara-Díaz, 2000, Hensher, 2001, Mackie et al., 2001).
6. The same holds true for parameters α and ϵ in the travel cost functions of households (Owen et al., 1992, Rus, 1990, Abdelwhab, 1998, Hensher, 1998, Hansen et al., 1996, Sonesson, 2001).
7. Net transfers N_r are assumed to be equal per capita within each country. The respective national value is the current account obtained from the social accounting system. Note that $\sum_r N_r = 0$ for the whole world. In the simulation experiments the N_r are held constant in real terms.

8. The parameters ψ^i appearing in equations (3.6) and (3.7) are arbitrary, just fixing units. Units for outputs can also be chosen freely for each sector. We choose units such that prices equal unity on average in the benchmark.
9. Finally, elasticities of substitution between varieties σ^i and transport cost mark-ups τ_{rs}^i must be chosen. These two issues are related and are discussed more deeply in the following, because these parameters are the most important ones for the outcome of simulations.

According to (3.13) and (3.14), τ_{sr} depends on transport distance g_{sr} and the parameters ζ , ω and δ_{mn} . In order to see how the other parameters appear in observed trade flows, insert ϕ from (3.14) into (3.13) and τ_{rs} from (3.13) into equation (3.7) describing trade flows. Equation (3.7) is then rewritten in gravity form (still omitting superscript i),

$$p_r t_{rs} = A_r B_s \exp(-\sigma \zeta g_{sr}^\omega) \delta_{mn}^{-\sigma}, \quad (3.15)$$

with

$$\begin{aligned} A_r &= X_r p_r^{1-\sigma}, \\ B_s &= D_s q_s^\sigma. \end{aligned}$$

In fact we do not have sufficient observation on interregional trade for directly estimating Equation 3.15. But let us assume for a moment we had such data. How to estimate 3.15? First, we have to specify δ_{nm} . Let $\log \delta_{nm}$ be linearly dependent on a set of explaining variables gathered in a vector z_{nm} , that means $\log \delta_{nm} = \pi' z_{nm}$ with parameter vector π ⁵. Inserting this into 3.15 and expanding yields

$$\begin{aligned} p_r t_{rs} &= A_r B_s \exp \left[-\sigma \zeta \omega \left(\frac{g_{sr}^\omega - 1}{\omega} \right) - \sigma \zeta + \pi' z_{nk} \right] \\ &= \exp \left[a_r + b_s - \rho g_{rs}^{(\omega)} + \pi' z_{mn} \right] \end{aligned} \quad (3.16)$$

with $a_r = \log A_r$, $b_s = \log B_s - \sigma \zeta$, and $\rho = \sigma \zeta \omega$. $g_{rs}^{(\omega)}$ denotes the Box-Cox-transform,

$$g_{rs}^{(\omega)} = \frac{g_{rs}^\omega - 1}{\omega}$$

According to Equation 3.16, the logs of observed flows, valued at fob-prices, are linear in a set of explaining variables, among them row and column dummies (with parameters a_r and b_s) and one Box-Cox-transformed variable. Note that, with $\omega = 0$ and $\omega = 1$ we obtain the power form and the exponential form of the distance function as special cases. According to our assumptions, however, we should obtain $0 < \omega < 1$. Now we can add a random disturbance to the RHS of 3.16 and fit it to the observations by choosing a_r , b_s , ρ , ω and π maximising the likelihood.

⁵ Remember that region r is in country m and region s in country n . $z_{mn} = 0$ for $m = n$.

As we in fact do not have the required data on a regional scale, we use international trade flows instead, assuming that Equation 3.16 is also valid for aggregated flows crossing the border. The regression then reads

$$\tilde{t}_{mn} = \exp \left[a_m + b_n - \rho \bar{g}_{mn}^{(\omega)} + \pi z_{mn} \right] + v_{mn}$$

$\tilde{t}_{mn}$ is the fob-value of trade from country m to country n . a_m and b_n represent fixed effects of the export and import country, respectively. v_{mn} is a random disturbance. $\bar{g}_{mn}$ is the weighted average distance from regions in country m to regions in country n , with regional GDPs taken as weights. Regarding z_{mn} , we try dummies for existence/non-existence of a common border, existence/non-existence of a common language, a dummy taking on a value of one if the respective flow crosses the former iron curtain and zero otherwise, and a few more dummies. Earlier studies of this kind show that very stable estimates with plausible parameters are obtained by these regressions (see Bröcker, 1999, 2000, 2001).

Once we have parameter estimates $\hat{\rho}$ and $\hat{\omega}$ we can take $\hat{\rho}/\hat{\omega}$ as an estimate of $\sigma\zeta$. Hence our transport cost function reads

$$\phi(g_{rs}) = \exp \left(\frac{\hat{\sigma}\zeta}{\sigma} g_{rs}^{(\omega)} \right). \quad (3.17)$$

How to calibrate σ ? A limit to the value of σ is given by the fact that, given $\phi(g_{rs})$ as in Equation 3.17, the transport cost intensity C , defined as the average ratio of transport costs⁶ to the value of trade, is decreasing in σ :

$$C = \frac{\sum_{rs} p_r \hat{t}_{rs} \left[\exp \left(\frac{\hat{\sigma}\zeta}{\sigma} g_{rs}^{(\omega)} \right) \right]}{\sum_{rs} p_r \hat{t}_{rs}} - 1$$

$\hat{t}_{rs}$ is the calibrated trade flow, which only depends on estimates of ρ and ω , not on σ .

With an independent estimate of the transport cost intensity C one can infer on σ . Estimates of transport costs and logistic costs, of which transport costs are a subset, can in fact be found in the literature. In a review of Weber (1987) logistic costs as a share of sales value vary between 12 percent and 22 percent, averaged over manufacturing industries. Mere transport costs, however, are close to 5 percent of sales value. Logistic costs include several components which are not related to distance and therefore should not be included in our estimate. On the other hand, our notion of distance costs includes components like costs of transferring information, which are clearly related to distance and not included in transport cost. Hence, distance cost intensity is probably in the order of 5 to 10 percent in manufacturing. More empirical information on this issue will come from the other project partners and from the SCENES data basis. Another independent information for guessing

⁶ Here we only talk about transport costs which depend on distance, not those depending on national borders.

σ is revealed by empirical studies on monopolistic price mark-ups (see Bröcker, 1995, 1998a), of which we will also make use.

3.4.5 Model Results to be Used in Project Assessment

The most important results for project assessment generated by comparative analyses using CGEurope are the monetary measures of regional welfare effects of the evaluated projects. They measure utility gains of regional households and translate them to monetary amounts by the concept of equivalent variation (EV). As noted above, one must not confuse these numbers with income changes. EV covers not just utility changes due to income changes.

Utility changes are generated by

- changes of factor prices, generating income changes (given constant factor stocks),
- changes of goods prices,
- changes of consumption goods diversity (reflected by changes of the composite goods prices q_s^i in our case), and
- changes of passenger travel times per unit of travel.

Other results which might be useful in an assessment outside the strict framework of CBA are:

- changes in passenger travel, by origin, destination and travel purpose;
- changes in interregional trade by sector, region of origin and region of destination, in nominal and real terms; it is worth mentioning that, unlike engineering models of travel flows, CGEurope measures real flows not in tons, but as values in constant benchmark prices (like the real GDP, for example);
- output changes, again in real and nominal terms;
- nominal and real income changes;
- factor price changes.

Estimates of real output and flow changes may also help in estimating environmental effects and adding monetarised environmental costs and benefits to the CBA results. These are not covered by the welfare measure generated in CGEurope.

Finally, note that CGEurope, according to present plans, does neither predict employment effects nor migration effects. Employment is held constant; labour demand adjusts to the fixed level of employment by flexible wages. The spatial distribution of the population is held fixed as well in the comparative static simulations.

3.5 The CGEurope Model Data Requirements

Generally there are three types of socio-economic data required by the CGEurope model: (i) *regional data*; (ii) *national accounts data*, and (iii) *transport cost data*. It requires information at a NUTS-2 level of sectoral activity of industries output, input, employment, transport costs and passenger travel, trade and other variables that allow meeting basic requirements of as-

assessment of spatial economic impacts of transport projects and policies (see Section 5 for a scenario overview). This section specifies in detail socio-economic data requirements of CGEurope. Transport data – passenger flows, transport costs, households income shares for travel and transport costs changes in turn of changes in policy scenarios – are part of the common database and, therefore are discussed last.

Given the degree of country and regional specificity and the numerous different policy objectives, which application of CGEurope will have in the present study, it is impossible to use the same sectoral and regional disaggregating for all countries (see Table 4.1 in Section 4). Therefore, a differentiation is made in the CGEurope model's disaggregating between three groups of countries: EU member and non-member countries and between the rest of the world regions (see Table 3.1 and Section 4).

Table 3.1. Details of CGEurope-specific data requirements by country group

	<i>Regional data</i>	<i>Country-level data</i>	<i>Transport costs</i>
<i>Compiler</i>	<i>CAU</i>	<i>RUG</i>	<i>IRPUD/TNO</i>
'EU'	N2/N3	N1	N3
CESE	N2	N1	N2
ROW	N1	N1	N1

N0 = NUTS-0 regions
 N1 = NUTS-1 regions
 N2 = NUTS-2 regions

'EU' = EU 15 + Switzerland and Norway
 CESE = Central, Eastern and Southern Europe
 ROW = Rest of the world (non-European countries)

The database of CGEurope will be organised in a matrix form as Social Accounting Matrix (SAM). SAM is a table consisting of rows and columns representing sectors of the economy. These sectors correspond to five main accounts in the regional economy: production activities, factors of production, institutions, capital, and the rest of the world (Isard et al., 1998). The number of rows in a SAM table must equal the number of columns; that is, each sector in the economy is identified with both a row and a column.

The SAM framework has an 'open architecture' and can be adjusted in the future to answer specific transport policy questions, if additional data become available. Depending on the specific transport policy area, certain accounts can be described through a larger number of detailed sectors in the future. For example, if transport policy makers are concerned about the distribution of income to different households based upon a change in government expenditures, then creating multiple household sectors based on income level can extend CGEurope. Likewise, CGEurope may define the detail of industry accounts in the model to better address policy questions. For example, it may be useful to disaggregate the transport sector into multiple transport commodities if the region is highly dependent on transport sector. Thus, the impacts of exogenous changes in final demand for any transport commodity can be compared to another transport commodity's impact on the local economy.

The CGEurope model will be benchmarked to the year 1997, which implies the current model's database has to be updated to the year 1997 and new variables introduced into the model have to be selected for this year.

3.5.1 Regional Data

For the purpose of assessing impacts of transport projects and policies on a regional level, a spatial CGE model requires information for each sector in each region, all inputs by sector and region of origin. In detail, these are those described in this Section – information on the national accounts data and the input-output coefficients; and additional information on the regional level, such as information on location of sectors (such as GDP/GVA by sector and region), regional factor prices and household income (spending). Ideally, sectoral output data would be collected for the CGEurope model implementation. Nonetheless, it is important to distinguish between desirability and feasibility. The existing information base and the time and resource constraints of Task 2.3 are key factors to be considered when deciding which data sources to use. If only GVA or employment data are available at regional level, it will still be possible to calibrate the CGEurope model for the benchmark year.

- *Gross Domestic (Regional) Product.* Gross Regional Product and sales are required to present industries' activities of the base year (1997). In a Social Accounting Matrix (SAM), output differs from sales because it includes changes in stocks of finished goods and work in progress and because of differences in measurement applicable to activities involving trade or intermediation. Gross domestic product and gross regional product estimates could be based e.g. on Eurostat's Regional Accounts by activity tables and could use as a first source data provided by the European Statistical Office in the New Cronos database. However, since activity detail is only provided at fairly aggregate levels in the New Cronos database (currently three NACE Rev. 1 activities - agriculture, services and industry), maximum use should be made of supplementary databases.
- *Value Added.* To reduce heterogeneity in the data, empirical implementation of the new version of CGEurope requires gross value added estimates of six industries at a regional level. Gross value added enters the database as the amount by which the value of the outputs produced exceeds the value of the intermediate inputs consumed. Although it is defined in terms of outputs and intermediate inputs, in a SAM value added also is equal to the sum of primary incomes generated in production (compensation of employees, profits, etc.). In some cases, depending on the particular data that are available, this equivalence could be exploited in deriving estimates of value added. The alternative of primary incomes generated in production could be chosen, for example, if data on intermediate consumption are lacking but information on the various incomes generated in production are available.
- *Employment.* In most of the statistical sources, employment is measured as the number of persons on the payrolls of in respective industries. In the common database, employment data would preferentially be converted to a 'full-time equivalent' (FTE) basis, in which part-time workers are counted according to the time worked (e.g., two workers on half-time schedules count the same as one full-time worker). Although FTE employment may provide a better measure for assessment of spatial economic impacts of transport projects and policies purposes, this measure is not as widely available as number of employees and is difficult to implement consistently. For these reasons, employment variable in the common database will be the number of persons employed in the base year. The number should be chosen representative of the base year (1997), in the absence of strong seasonal and other fluctuations in employment and should be measured as of a point in time, (the end of the year), following national practices.

- *Regional final demand.* As there are only scarce final demand data on a regional level it will be necessary to estimate the missing final demand variables by starting from regional GDP and adding net inflows of income and net capital inflows, which are based on plausible assumptions, such as equalising the public budget and savings-investment closure (neither public sector final demand nor the investment demand are treated explicitly in the model).

Because of missing regional data, it would be impossible to establish a full SAM on a regional level by using only the information available and to solve the balancing conditions of the SAM since there would be too many unknowns and too few equations. To deal with regional data scarcity problems in spatial CGE models, several approaches have been used in the literature. One approach is that the modeller, based on his or her knowledge of the local economy, may adjust cell values within the SAM to equate row and column totals. A non-economic approach that has been used is the RAS technique⁷ (Isard et al., 1998, Miller and Blair, 1985). Another technique that combines features of the RAS method with additional information provided from the SAM is the cross entropy approach (Robinson et al., 2001). Each of these approaches assist the modeller in using scarce data while at the same time achieving accounting balance within the SAM.

CGEurope does not follow these approaches, as the theory behind these remain obscure. Instead of using such ‘data generating’ procedures, CGEurope assumes that production technologies of firms and household preferences do not depend on location. Therefore, detailed social accounting information is only required on a national scale. These are described in the following section.

3.5.2 National Accounting Data

CGEurope does not have the entire information required for building a regional SAM. Three kinds of data are *not available* on a regional scale: (i) national accounting identities, (ii) input-output coefficients, and (iii) inter-regional trade data. The national accounts data described in this section combines information on national accounts for each country, including input-output information with institutional flows of goods and services and linking countries through trade of goods and services.

National accounts comprise main accounting identities of each economy and are based on a principle of double-entry bookkeeping, which is required in the form of payments and receipts also by the SAM. National accounts data serve key features for constructing and balancing the SAM. As a primary source of national accounts data, CGEurope could use e.g. the OECD's National Accounts, which provide, in addition to main aggregates, estimates broken down by kind of activity for gross value added, components of value added, gross fixed capital formation and employment. OECD's National Accounts of central and eastern Europe could be also of interest for the common database of Work Package 2.2 because they contain information of GDP by the income, expenditure and production approaches for the 10 CEE countries: Bul-

⁷ The RAS approach is a mathematical procedure, in which a new coefficient matrix is generated by solving an optimisation problem subject to given row and column margins, represented by the totals of intermediate output and intermediate purchases.

garia, the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, the Republic of Slovenia, Romania and the Slovak Republic.

Besides economic activity data, input output coefficients are probably the most important in transport project and policy evaluation. The statistical information on activity tables and uses characterise economic linkages of a base year. Industry use and make tables present differences in the sectoral factor productivity between countries. In the context of CGEurope, input-output coefficients are also required for assigning interregional and international flows to sources and destinations, by combining input-output information with a gravity approach, which is derived from microeconomic foundation (see Sections 3.3 and 3.4). For most of the countries, no accounting data are available showing input-output flows by sector and region. The methodological focus of CGEurope is therefore on designing a multiregional and multi-sectoral SAM applicable to poor data environment, such as CESE countries. Even national input-output information shows no complete social accounting matrix and is published only with long time lags, so that updating and filling of data gaps by plausible assumptions cannot be avoided. To overcome this shortcoming in the database, CGEurope has restricted the number of parameters such that all parameters except elasticities can be calibrated by national input-output data and by regional data showing no more than regional employment and other activity indicators by sector and regional factor prices (see Section 3.3).

Eurostat's Input-Output tables (IOT) could provide first estimates on inter-industrial transactions (domestic and import input-output matrices) for the SAM. The last published version comprises tables for 1995 for the 15 EU member countries with a CLIO 25 industry classification, which is not consistent with the new version of CGEurope yet. The ongoing work at RUG to provide updated tables based on an ISIC Rev. 3 industry classification consistent with common industry classification asks for compiling activity tables and uses a six sector level of industry (in national SIC) for as many CGEurope countries as possible. For the rest of the countries, where input-output data is missing, proxies have to be used. See below for a description of the six economic sectors used.

A further important variable in the structure of the CGEurope model is the international and interregional exchange of goods and services. This importance is given by the fact that the purposed analysis will concentrate on the consequences that transport projects and policies have in the region in question as well as how much of these consequences flow outside the region. Therefore, it is required to estimate the trade flows. For the calibration only trade information on the international level is required, while the interregional estimates result from the calibration.

Information on exports and imports flows of goods by industry could be extracted, e.g., from the UC Davis World Trade Analyser (WTA) database. The WTA provides a breakdown of trade flows by manufacturing industry between more than one hundred countries and a selection of partner countries as well as larger geographical entities. The latest WTA version covers the year 1997 and uses an ISIC Rev. 2 industry classification, which would be consistent with CGEurope.

All national data from various sources have to be compiled into one database. This work will be performed in the task WP 2.3 and includes among others gathering of data, aggregating to the common regional and sectoral level, adjusting the overlapping values from different sources and checking the CGEurope database for consistency.

3.5.3 Economic Sectors

An important feature of the common database is the use of a standard industry classification to facilitate comparisons between the results of the SASI and CGEurope models. In addition to data considerations of availability, the following criteria have been applied to define economic sectors:

- The sector is important to the national economy and in particular in its contribution to national GDP;
- The sector has been, or might become, the subject of changes in economic rules induced by transport-related policies;
- The sector is one with significant transport flows in both volume and financial terms and is experiencing changes in transport flows;
- The sector is one where one might expect, a priori, that there are important substitution effects attributable to transport-related policies.

The refined industry classification is based on ISIC Rev. 3 and will include six instead of two sectors. This preliminary sectors classification is due to changes with regard to data availability. Final decision about industry classification will be met in task 2.2. Currently it includes the six following activities:

- *Manufactured products* (ferrous and non-ferrous ores and metals, other than radio-active, non-metallic minerals and mineral products, chemical products, metal products, except machinery and transport equipment, agricultural and industrial machinery, office and data processing machines, precision and optical instruments, electrical goods, transport equipment, food, beverages, tobacco, textiles and clothing, leather and footwear, paper and printing products, rubber and plastic products, other manufactured products);
- *Market services* (recovery and repair services, wholesale and retail trade services, lodging and catering services, inland transport services, maritime and air transport services, auxiliary transport services, communication services, services of credit and insurance institutions, and other market services);
- *Agriculture, forestry and fishery products* (agriculture, hunting and related services, forestry, logging and related services, fishing, operation of fish hatcheries and fish farms);
- *Fuel and power products* (electricity, gas and water supply and collection, purification and distribution of water);
- *Building and construction* (construction services covers work performed on construction projects and installation by employees of an enterprise in locations outside the territory of an enterprise);
- *Non-market services* (public administration and defence, compulsory social security, education, health and social work, and other non-tradable community, social and personal service activities).

This classification is compatible with the NACE Rev. 1 classification used by the EU member countries. The industry classification is designed to provide CGEurope and SASI with enough

sectoral detail to focus on transport- and/or shipping-intensive industries while taking into consideration general data availability across countries (based on recent experience). Sectoral information coverage for each country depends on: (i) whether national statistical offices compile the information by industrial activity in the context of regional accounts; (ii) the extent of updating made by national statistical offices after the recent widespread revisions of national accounts; (iii) finally, availability of regions socio-economic data at NUTS 2 / NUTS 3 level (for considered sectors). Since activity detail is only provided at fairly aggregate levels for most of the Central and Eastern European countries the Joint Spatial Economic Database industry classification includes an alternative aggregate for use when modelling transport investments and policy, across non-member countries. Currently this aggregate classification in the Central and Eastern European countries include the following three activities:

- 'Agriculture' covering A and B a NACE industry categories;
- 'Manufacturing' covering C, D and E a NACE industry categories; and
- 'Services' covering the rest of the industry (F to P) a NACE industry categories.

The industry classification may expand in the future in response to changes in data availability and in project assessment requirements. Empirical implementation of the new manufacturing industry groups based on 'transport intensity' involves then ranking industries according to indicators based on the use of acquired transport. The base year will be 1997, which is the latest year for which a comprehensive information on socio-economic variables is available.

3.5.4 Transport Costs

Transport costs for goods and services are – beyond the spatial distribution of supply and demand – the main determinants of trade flows in the CGEurope model. Transport policy scenarios eventually influence variables of interest like GDP and welfare through changing transport costs.

Transport costs need not only be quantified for goods, but also for passengers. Passenger transport costs are another determinant of trade in goods and services, because trade relations require face-to-face contacts, and they are the main determinant of private passenger flows (beyond the spatial distribution of households and of destination opportunities).

For estimating travel and transport costs information is required about time costs measured as shares of goods values, about the levels of other kinds of transport costs, and about the reaction of these costs to transport distances. Furthermore, for passenger travel one needs information on the relation between monetary travel costs to travel distances as well as on willingness to pay for time savings (see the following section).

Travel cost and trade flows estimates have already been compiled in the SCENES project. These data should flow into the CGEurope database.

3.5.5 Passenger Transport Flows

Long distance mobility information is indispensable for the empirical implementation of the enhanced CGEurope and SASI models to assist infrastructure planning through the TEN programmes and transport policy measures. The creation of passenger travel and freight transport data infrastructure requires setting up a system on inter-regional travel behaviour in all EU as

well as the CESE countries considered in the models. No full information tables will be available but only aggregate and incomplete information, which will be used as additional information for model calibration.

Within CGEurope, not only effects of transport cost changes for goods and services, but also for private long distance travel, which could also imply inter-regional travel, will be evaluated. For this the CGEurope model needs two kinds of information, namely travel expenditure shares of private households, and a benchmark matrix of interregional passenger flows. These two data sets are expected to be available while only limited additional data required for updating and for adjustment to one system of regions.

As a 'primary source', origin-destination matrices for passenger movements at the NUTS-1 level for all Europe with a predefined level of confidence for the estimations produced should be used from the databases compiled in SCENES and in the ETIS projects BRIDGES and CONCERTO. Furthermore, a number of member states have long-standing surveys, in particular of freight traffic and daily private travel, combined with extensive counting programmes on their road, rail, air, inland-waterway and maritime networks and nodes. Where possible, the passenger travel data from the SCENES and ETIS projects should be coupled and integrated with national counts survey data.

4 Regions and Networks

This section presents the framework for the Common Spatial Database to be used by both the extended SASI and the CGEurope models: the system of regions and the network database. Here only the basic principles for developing the Common Database are presented. The Common Database will be presented in detail next.

4.1 System of Regions

The system of regions defined is based on level three of the Nomenclature of Territorial Units for Statistics (NUTS) for EU member states (Eurostat, 1999a) and equivalent regions for the candidate countries (Eurostat, 1999b). Because for Poland negotiation on NUTS 3 level regions are still pending, NUTS 2 level regions are used for the moment but will be substituted as soon as NUTS 3 level regions are available. For the other European countries, only a limited number of regions is defined (Table 4.1). With the exception of Belarus (6 regions), Switzerland (26), Norway (19), Russia (28) and Ukraine (3), all other countries are not further subdivided (see Figure 4.1).

The 1,083 regions defined for the EU member states are the so-called 'internal' regions of the model. 162 regions located in candidate countries are designated as 'candidate' regions, whereas 91 regions are 'external' regions for the rest of Europe, and five regions representing the 'rest of the world'. The five regions representing the rest of the world are only used as origins and destinations of freight flows, but economic performance indicators are not calculated for them. Altogether, 1,341 regions are defined. Table A-1 in the Annex gives a full description of these regions including their main economic centres.

4.2 Trans-European Transport Networks

The spatial dimension of the system of regions is established by their connection via networks. The economic centres of the regions are connected to the network by so-called access links. Road, rail and air networks are considered. The 'strategic' road and rail networks defined are subsets of the pan-European road and rail network database developed by IRPUD (2001), comprising the trans-European road links specified in Decision 1692/96/EC of the European Parliament and of the Council (European Communities, 1996) and specified in the TEN implementation report (European Commission, 1998), the TINA networks as identified by the TINA Secretariat (1999), the Helsinki Corridors as well as selected additional links in eastern Europe and other links to guarantee connectivity of NUTS-3 level regions and centroids. The 'strategic' air network is based on the TEN and TINA airports and other important airports in the remaining countries considered and contains all flights between these airports.

The networks will be used to calculate regional accessibility. For that the historical and future developments of the networks are required as input information. This development of the networks over time is reflected in intervals of five years in the database, i.e. the established network database contains information for all modes for the years 1981, 1986, 1991, 1996, 2001, 2006, 2011 and 2016. The way the historical and future dimensions of the network are established in the GIS database is described in detail in the framework of the SASI project (Fürst *et al.*, 1999, 30).

Table 4.1. Number of regions

Region	Country	Number of regions
EU member states	Österreich	35
	Belgique/Belgie	43
	Deutschland	441
	Danmark	15
	Espania	48
	Suomi/Finnland	20
	France	96
	Ellada	51
	Ireland	8
	Italia	103
	Luxembourg	1
	Nederland	40
	Portugal	28
	Sverige	21
	United Kingdom	133
<i>Total EU member states</i>		<i>1,083</i>
EU candidate countries	Balgarij	28
	Cyprus	1
	Cesko	14
	Eesti	5
	Magyarország	20
	Lietuva	10
	Latvija	5
	Malta	1
	Polska	16
	Romania	42
	Slovenija	12
	Slovensko	8
<i>Total EU candidate countries</i>		<i>162</i>
Rest of Europe	Shqiperia	1
	Bosna I Hercegovina	1
	Belarus	6
	Schweiz	26
	Hrvatska	1
	Island	1
	Liechtenstein	1
	Moldova	1
	Republica Makedonija	1
	Norge	19
	Rossija	28
	Türkiye	1
	Ukraina	3
	Jugoslavija	1
<i>Total rest of Europe</i>		<i>91</i>
Rest of world	North-America	1
	Latin America	1
	Africa	1
	Middle East	1
	Asia, Australia	1
<i>Total rest of world</i>		<i>5</i>
<i>Total all regions</i>		<i>1,341</i>

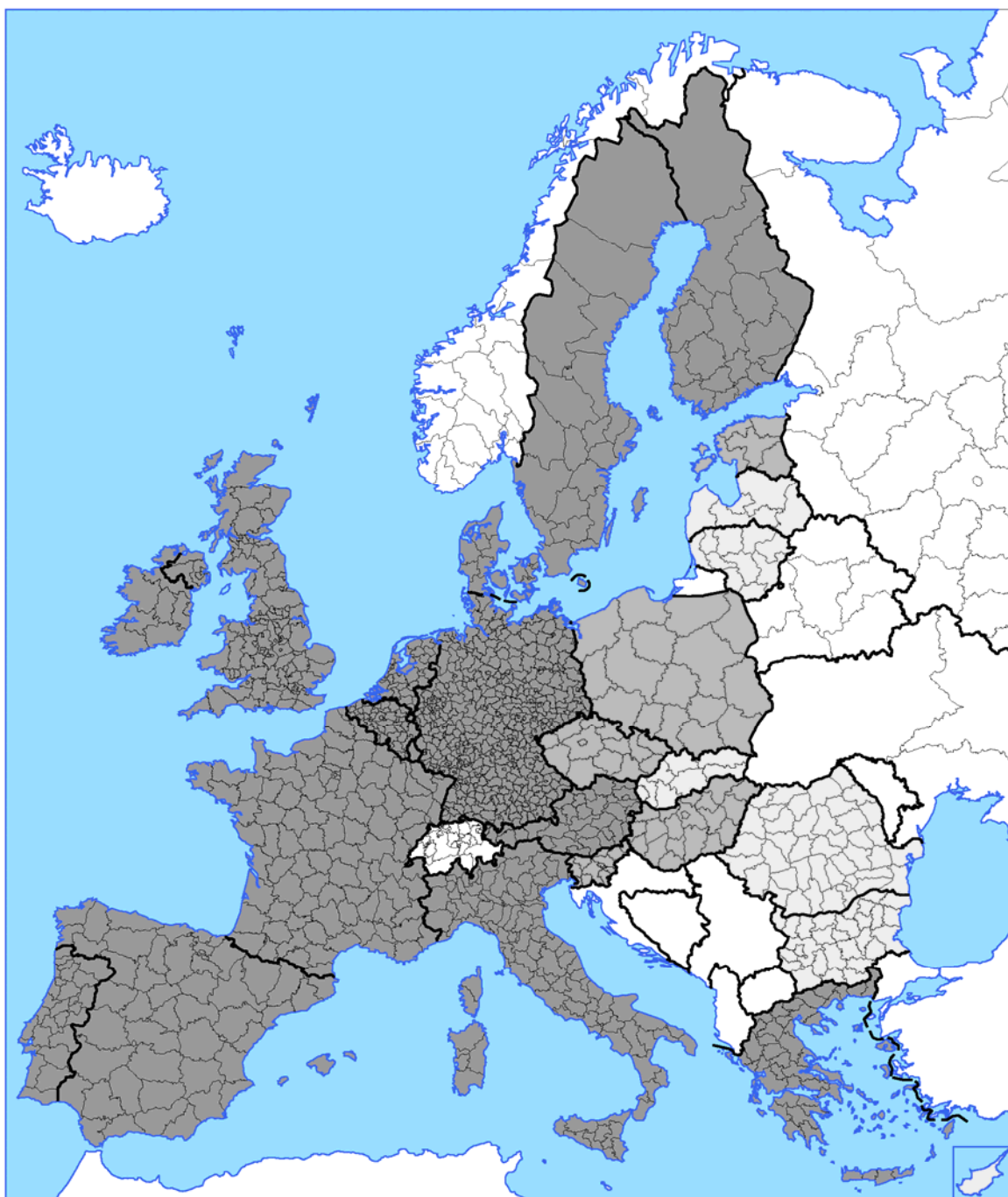


Figure 4.1. The system of regions

Road network

The strategic road network contains all existing and planned motorways, all dual-carriageway roads and other expressways, all E-roads and main international traffic arteries identified by UN (1995), the most important national roads and car ferries, the Eurotunnel and additional motorail links (road/rail interchange points for Alps crossing), as well as additional minor or secondary roads to guarantee connectivity of NUTS 3 region centroids (Figure 4.2).

The road network database contains information on the type of road ('link category'), inclusion in the TEN and TINA programmes, time penalties in agglomeration areas due to congestion and in hilly areas due to slope gradients, ferry timetable travel times, road tolls, national speed limits and border delays.

Link categories of past networks are compiled from Shell (1981; 1992), ADAC (1987; 1991), Reise- und Verkehrsverlag (1987) and Michelin (1992a; 1992b). Link categories of future networks are taken from the TEN implementation report. National speed limits are derived from ADAC (2000), and assumptions on border waiting times are based on IRU (1998) (see also Fürst *et al.*, 1999; Schürmann and Talaat, 2000a; 200b). Figure 4.3 gives a representation of the future road network evolution until the year 2016 according to the envisaged completion and opening years of the road projects.

Rail network

The strategic rail network contains all existing and planned high speed lines, upgraded high speed lines and the most important conventional lines as well as some rail ferry and other minor or secondary lines to guarantee connectivity of NUTS 3 regions centroids (Figure 4.4).

The rail network database contains information on the type of link ('link category'), inclusion in the TEN and TINA programmes and timetable travel times.

For the past rail networks, it was first checked which railway line already existed in 1981, 1986 and 1991 and which not. For example, most of the current links existed already in 1981 with the exception of the new high-speed lines (Fürst *et al.*, 1999, 35). In order to have the connectivity of the current high-speed lines in the 1981 network, corresponding conventional links are introduced in the 1981 strategic rail network. The new high-speed links are introduced into the strategic networks of 1986, 1991 or 1996 according to their opening year. Moreover, for the remaining lines, assumptions have been made for the general increase of the 1996 timetable travel times due to technical improvements in signalling techniques.

The TETN implementation report contains information on planned new (high speed or conventional) lines or planned upgraded lines (see Figure 4.5). This information is used to make assumptions for speed and travel time changes on a country-by-country basis with respect to the new link categories. In some cases published future travel times for railway sections are used. If no upgrading is planned for a link, a modest acceleration of ten percent is assumed which reflects improvements in signalling systems, carriage technology and railway construction.

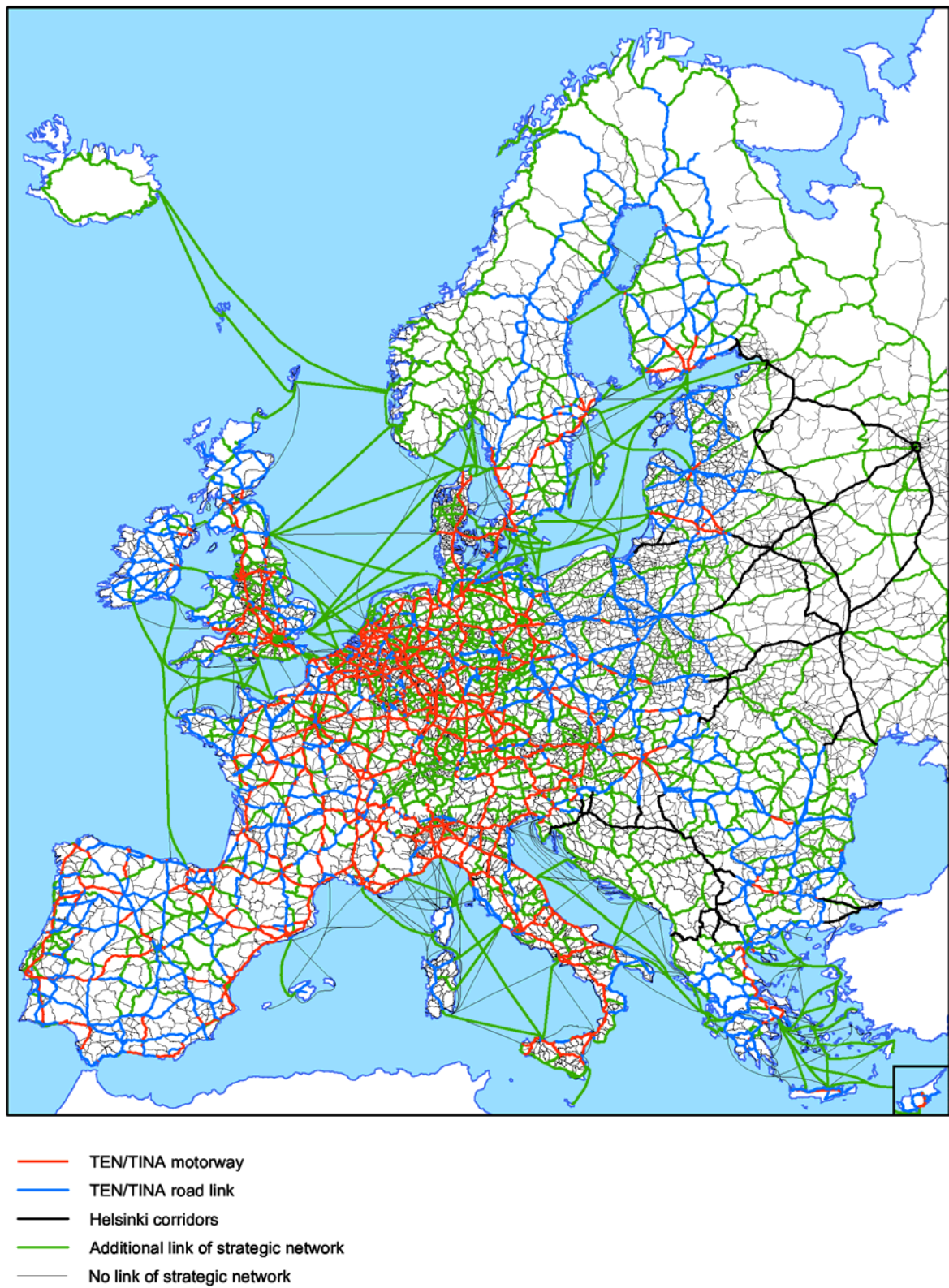


Figure 4.2. The strategic road network in 2001



Figure 4.3. Road projects according to TEN/TINA outline plans



Figure 4.4. The rail network by link category in 1996



Figure 4.5. Railway projects according to TEN/TINA outline plans

Air network

The generation of the strategic air network had to be different from the generation of the road and rail networks (Fürst *et al.*, 1999, 37). This is because air networks do not consist of physical link infrastructure. The only physical infrastructure are the airports. Therefore, the generation of the strategic air network started with the definition of airports of strategic interest.

The airports forming the base of the strategic air network are all airports contained in the TEN and TINA programme. In addition, important airports in eastern Europe and other non-EU countries are included to guarantee connectivity of these regions (Figure 4.6).

The criterion for an airport to be a node in the air network is that it has at least one regular daily flight. Eight smaller airports (according to the TEN nomenclature so-called 'Regional and Accessibility Points') have only charter flights or flights on demand and have been excluded from the strategic network.

All in all there are about 330 airports establishing nodes of the air network. All regular flight connections between these airports form the 1996 air network. The information has been extracted from the *Air Traffic Databank* produced and maintained by Mkmetric (1998). The air network contains only non-stop relations between two airports. This means, for example, a flight from Madrid to Berlin via Frankfurt is divided into two flights, the first one from Madrid to Frankfurt and the second one from Frankfurt to Berlin. Furthermore, outward and return flights are stored as two separate relations. In total, there are 4,156 relations stored in the database. Charter flights, non-regular flights or tourist flights are not included.

Average travel times (expressed as scheduled flight times as average travel time calculated over all flights over all wind exposures over all kind of planes), terminal times and the number of flights (expressed as a frequency index over the year) are associated with each flight connection.

The creation of past air networks is a difficult task. There is no source available which gives air networks for the past for entire Europe. Therefore simple assumptions had to be made about the air networks for 1981, 1986 and 1991. The basic assumption is that regional airports played a minor or no role at all in the beginning of the 1980s. This was reflected by adding a time penalty on 1996 air travel times for flights going from or to regional airports: the time penalty was 30 percent for 1981, 20 percent for 1986 and 10 percent for 1991. Moreover, some of the regional airports and so the flight relations to or from these airports were dropped from the past networks.

The generation of the future air network is a difficult task as well. Because the basic characteristic of the air network is that all airlines design their own flight connection system on own responsibilities, there are no official plans or even planning authorities for the development of the air network. Given that and the focus of the project on changing rail and road infrastructures, the future air networks will be the same as the current air network, i.e. no changes will be implemented, except of different assumptions on terminal times.

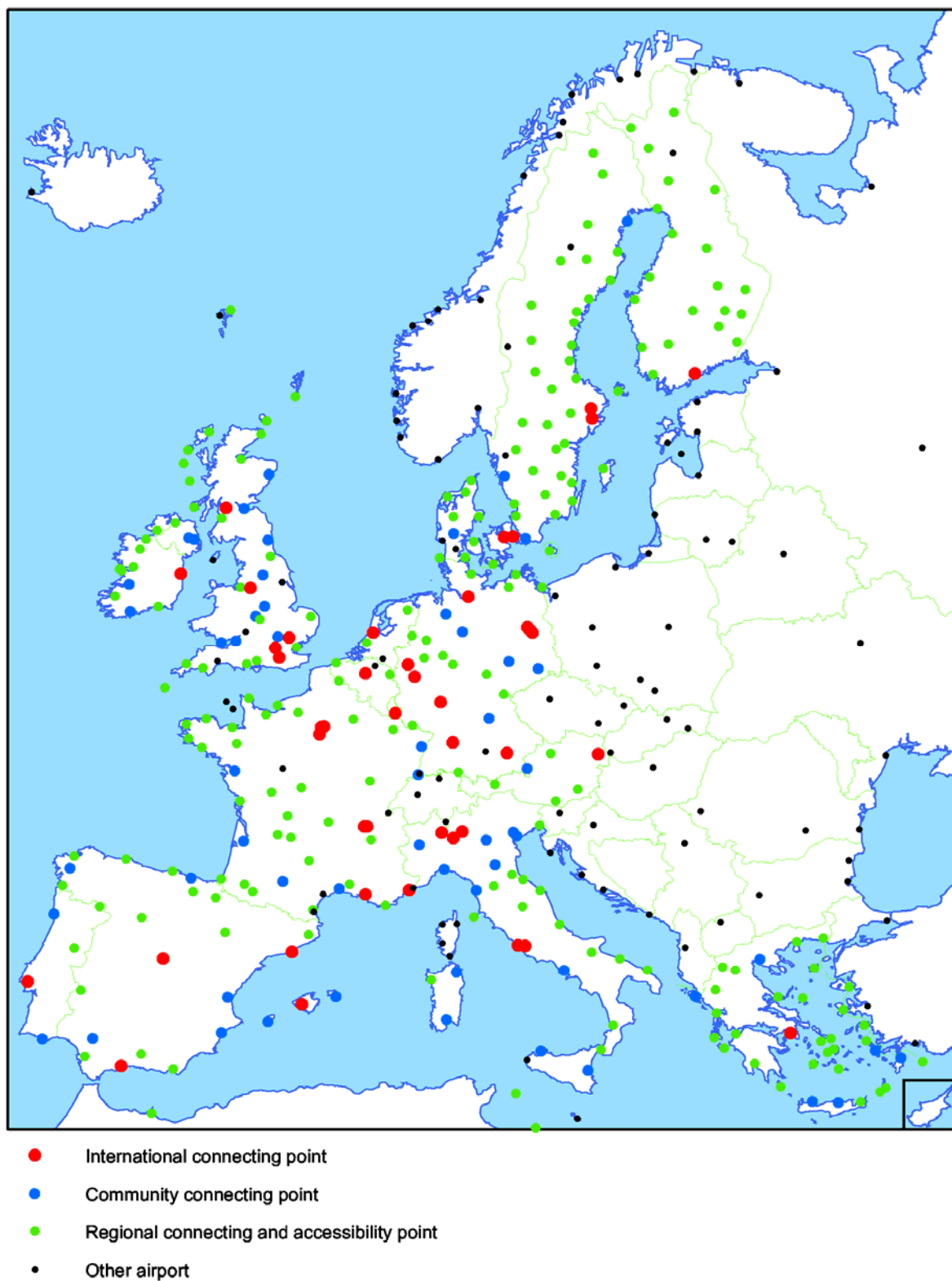


Figure 4.6. Airports by international importance

The aforementioned brief descriptions of the road, rail and air networks could only be a first attempt to a broader and more comprehensive description of the database that will be done subsequently. There, a detailed description of the network and socio-economic database will be given along with a description of tools and macros to extract and pre-process the database to fit input requirements of the two models.

5 Possible Scenarios

A scenario in the present study will be a time-sequenced programme consisting of a combination of policies in the fields of transport, economy and migration. In technical terms, a scenario will be any combination of assumptions about the development of the trans-European network infrastructure, European/national transport policies, total European GDP, European/national transfer policies, total European migration and European/national migration policies.

Two fundamental groups of policies can be distinguished scenario applications, either considered separately or together in a certain scenario. The first group of policies affect the European transport infrastructure and its use. The second group change the socio-economic macro-trends assumed for the transport scenarios. Both groups of scenarios are outlined in the following paragraphs.

5.1 Transport Infrastructure Scenarios

The group of transport infrastructure scenarios can be further subdivided into *Network Scenarios* and *European/national transport policies*.

Network scenarios

Networks scenarios consist of assumptions about the development of the trans-European transport networks. These assumptions have the form of backcasts of the road, rail and air networks representing their evolution between 1981 and 1996 as well as forecasts of their development between 1996 and 2021, both in five-year increments (see Fürst *et al.*, 2000).

The *Do-Nothing Scenario* is defined as a development in which no network changes are implemented after 1996. Other scenarios, in which specific subsets of the TEN are implemented are:

- *TEN Scenario*: all links specified in the TEN/TINA masterplans are implemented.
- *Rail TEN Scenario*: only rail TEN links are implemented.
- *Road TEN Scenario*: only road TEN links are implemented.
- *Priority Projects Scenario*: all TEN priority projects are implemented
- *High-Speed Rail Scenario*: only high-speed rail links are implemented.

In addition, any individual TEN/TINA or non-TEN/TINA project, or any combination of road, rail, air or inland waterway projects can be examined by simulating a scenario with and one without the project or combination of projects.

European/national transport policies

Transport policy scenarios consist of assumptions about policy decisions affecting the use of the trans-European and other transport networks. Transport policies that can be examined with the extended models under development will be:

- changes in national speed limits

- changes in local speed limits in agglomerations or on specified road links
- changes in fuel prices
- changes in rail fares
- changes in rail travel times
- changes in air travel times
- changes in the number of daily flight connections
- changes in toll charges
- changes in car ownership/purchase taxes
- changes in ferry fares
- changes in border waiting times and cultural barriers
- changes in waiting times at ferry ports
- changes in waiting times at road-rail interchanges
- changes in statutory rest periods for drivers

5.2 Socio-Economic Scenarios

In addition to the transport scenarios, the socio-economic macro trends assumed for the transport scenarios can be changed. The most important macro trends considered refer to the overall development of the European economy and to the future development of immigration into the European Union (see Section 2.4.1).

European GDP

European GDP scenarios consist of assumptions about the future performance of the European economy as a whole. These assumptions have the form of observed values of GDP for each economic sector for the European Union as a whole and for the non-EU countries considered for the years 1981 to 1997 and of forecasts of the same for the years 1998 to 2021. In the scenarios assuming an enlargement of the European Union by the present candidate countries, the European GDP assumptions refer to the enlarged EU and the remaining countries, respectively.

European/national transfer policies

Transfer policy scenarios consist of assumptions about transfer payments by the European Union via the Structural Funds and the Common Agricultural Policy or by national governments to assist specific regions. These assumptions have the form of annual transfers received by any of the regions in the European Union during the period 1981 to 1997 and forecasts of the same for the period 1998 to 2021. These data only need to be provided for those regions that actually received aid in the past or are assumed to receive aid in the future.

European migration

European migration scenarios consist of assumptions about immigration and outmigration across Europe's borders. These assumptions have the form of total observed annual immigration from the non-EU countries to the European Union and total annual outmigration to these

countries from the European Union for the years 1981 to 1997 and of forecasts of the same for the years 1998 to 2021.

In the scenarios assuming an enlargement of the European Union by the present candidate countries, the European GDP assumptions refer to the enlarged EU and the remaining countries, respectively.

European/national migration policies

Migration policy scenarios consist of assumptions about immigration policies by European countries. Given the expected rapid population growth and lack of economic opportunity in many origin countries, total European immigration will be largely a function of migration policy decisions by national governments. These assumptions have the form of upper limits for annual immigration from non-EU countries to the countries of the European Union for the years 1981 to 1997 and of forecasts of the same for the years 1998 to 2021.

5.3 Scenario Applications

As indicated above, the above transport and socio-economic scenarios can be simulated each separately or in any reasonable combination. However, the above paragraphs outline the range of scenario applications that can be simulated with the CGEurope and SASI models and should be seen as first suggestions for possible applications. The actual number of scenarios applied, and their definition and selection will be decided on in co-operation with the other cluster partners.

6 Conclusions

The present study *Methodology for the Assessment of Spatial Economic Impacts of Transport Projects and Policies* sets up a methodological framework for the assessment of spatial economic impacts of transport projects and policies, by describing the extension and refinement of the two already existing Europe-level regional economic models, SASI and CGEurope. Furthermore, the study defines the system of regions, the sectoral categorisation, gives an overview about the models' requirements to the common data basis as well as gives first hints on baseline and alternative future year scenarios to be applied.

The existing SASI model will be updated and extended in several dimensions relating to model theory, model data and model technique (i) New ideas from growth theory as well as new evidence on firm location will be reviewed and transformed into operational indicators of locational advantage and disadvantage and incorporated into the econometric approach by forecasting rates of change rather than levels, modelling productivity endogenously, applying new indicators of accessibility incorporating wage levels and/or production costs and by modelling migration flows instead of net migration. In addition, efforts will be made to make the model more policy-relevant by extending the range of policies that can be simulated and expanding the cohesion indicators used. (ii) The model will be spatially more disaggregate and include the candidate countries in eastern Europe and use more recent and additional regional and transport network data. (iii) The model software will be extended and made more user-friendly and include more visual output.

Compared to the previous version of CGEurope the new version to be implemented in the present study has been enhanced in several ways: (i) The previous version had only two sectors (tradable and non-tradable), while the new one differentiates between six sectors, including one sector producing the transport service using factors and intermediate inputs; (ii) The previous version took only transport costs in inter-regional trade into account, while the new one also includes costs of private passenger travel; (iii) The new version of CGEurope models the use of resources for transport in a more sophisticated way than the previous one by including explicitly an activity producing the transport service; and finally (iv) The transport network from which the cost measurement is derived is much more refined, based on the networks developed within SASI, SCENES and ETIS.

Links between the SASI and CGEurope models have been established for a combined analysis of the transport projects and policies using these two models. The SASI model will provide transport cost and cost changes information for input into the CGEurope model, while the CGEurope model will provide socio-economic data so that the SASI model will have demand data that is consistent with transport cost accounting. The enhanced regression model can now be used to calculate the transport impacts of the network changes. The changes in transport costs will be incorporated into the multiregional general equilibrium framework. Reduced costs of intermediate goods will result in changes to final goods prices and hence demand. This will then change sectoral output and employment, investment. The changes in prices will also cause changes in the patterns of household expenditure and firms activity, which will be used to identify the distributional effects across regional households.

The structure of the enhanced SASI and CGEurope models will enable the indirect spatial economic impacts to be analysed in several different ways. Firstly, the impacts on the different industrial and service sectors can be identified, both within the EU member states and across the non-member countries. This will enable the impacts on European industrial com-

petitiveness to be assessed. Furthermore, the changes in employment and economic activity for each region will enable the distribution of indirect impacts and changes in growth prospects across the different regions to be determined. Estimated welfare effects by region will provide information for the socio-economic distributional analysis.

At the operational level, common data formats, for regions and for the representation of the industries have been defined. Both models use a common industry and regions classification in order to facilitate input data exchange and comparisons between SASI and CGEurope models' results. The sectors and regions definition has brought in line models' data requirements with the data availability.

Models' data requirements are enormous. They range from socio-economic indicators of regional GDP by sector, employment, input and output coefficients to behavioural parameters for CGEurope calibration. Some of the required statistical information is already available from the existing databases. Other data, such as regional GDP by sector and employment shares can be extracted from national and international statistical offices.

Transport data inputs such as data on passenger flows, households' travel expenditures, and information for converting travel times and travel lengths along shortest routes through the networks into travel costs have to come from the databases compiled in SCENES, TIPMAC and in the ETIS projects BRIDGES and CONCERT.

7 References

- Abdelwhab, W (1998): Elasticities of mode choice probabilities and market elasticities of demand: Evidence from a simultaneous mode choice/shipment-size freight transport model. *Transportation Research E* 34, 257-266.
- Allgemeiner Deutscher Automobil Club (ADAC) (1987): *Der große ADAC Europa Atlas. Deutschland und Europa*. München: Mairs Geographischer Verlag.
- Allgemeiner Deutscher Automobil Club (ADAC) (1991): *ADAC Straßenatlas 1991. Deutschland mit Europateil*. Berlin, Gütersloh, München, Stuttgart: Reise- und Verkehrsverlag.
- Allgemeiner Deutscher Automobil Club (ADAC) (2000): *Tempolimits*. <http://www.adac.de>.
- Aschauer, D.A. (1989): 'Is public expenditure productive?' *Journal of Monetary Economics* 23, 2, 177-200.
- Aschauer, D.A. (1993): 'Public capital and economic growth'. In The Jerome Levy Economic Institute (Ed.): *Public Policy Brief*. Annandale-on-Hudson: Bard College, 9-30.
- Batten, D., Boyce, D.E. (1986): Spatial interaction, transportation, and interregional commodity flow models," in Nijkamp, P. (ed.): *Handbook of Regional and Urban Economics*, Vol. I, Amsterdam: Elsevier Science, 357-406.
- Biehl, D. (Ed.) (1986): *The Contribution of Infrastructure to Regional Development*. Final Paper of the Infrastructure Studies Group to the Commission of the European Communities. Luxembourg: Office for Official Publications of the European Communities.
- Biehl, D. (1991): 'The role of infrastructure in regional development'. In Vickerman, R.W. (Ed.): *Infrastructure and Regional Development*. European Research in Regional Science 1. London: Pion, 9-35.
- Blum, U. (1982): 'Effects of transportation investments on regional growth: a theoretical and empirical investigation'. *Papers of the Regional Science Association* 49, 169-184.
- Bröcker, J. (1984): How do international trade barriers affect interregional trade? In: Andersson, Å. Isard, W., Puu, T. (eds.): *Regional and Industrial Development. Theories, Models and Empirical Evidence*. Amsterdam: North Holland, 219-239.
- Bröcker, J. (1995): Chamberlinian spatial CGE modelling: a theoretical framework. *Economic Systems Research* 7, 137-149.
- Bröcker, J. (1996): Measuring accessibility in a general equilibrium framework. Paper presented at the 5th World Congress of the RSAI, 2-6 May 1996.
- Bröcker, J. (1998a). Operational spatial computable general equilibrium modelling. *Annals of Regional Science* 32, 367-387.

Bröcker, J. (1998b): *Spatial Effects of Transeuropean Networks: Preliminary Results from a Spatial Computable General Equilibrium Analysis*. Diskussionsbeiträge aus dem Institut für Wirtschaft und Verkehr der Technischen Universität Dresden .4/98, Dresden: TU Dresden

Bröcker, J. (1999): The economic impacts of the TEN projects on the Øresund region: an equilibrium modelling approach. Paper presented at the International Seminar on Regional Development Impacts of the Øresund Bridge, Copenhagen, November 28, 1999.

Bröcker, J. (2000): Trans-European effects of "Trans-European Networks": results from a spatial CGE analysis. In: Bolle, F., Carlberg, M. (eds.): *Advances in Behavioral Economics*. Heidelberg: Physica, 141-159.

Bröcker, J. (2001): Spatial effects of European transport policy: a CGE approach. In: G.J.D.Hewing, M.Sonis, D.Boyce, editors. *Trade, Networks and Hierarchies - Modelling Regional and Interregional Economies*. , pages 11-28, Springer, Berlin, Heidelberg, New York.

Bröcker, J., Peschel, K. (1988): 'Trade'. In Molle, W., Cappelin, R. (Eds.): *Regional Impact of Community Policies in Europe*. Aldershot: Avebury.

Buhr, W. (1975): *Die Rolle der materiellen Infrastruktur im regionalen Wirtschaftswachstum*. Berlin: Duncker & Humblot.

Bundesministerium für Verkehr (BMV) (1993): Gesamtwirtschaftliche Bewertung von Verkehrsinvestitionen. Schriftenreihe des BMV, Heft 72, Bonn.

European Commission (1998): *Trans-European Transportation Network. Paper on the Implementation of the Guidelines. Basic Data on the Networks*. Paper to the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions on the implementation of the guidelines for the development of the trans-European transport network (Decision 1692/96/EC).

European Communities (1996): Decision No. 1692/96/CE of the European Parliament and of the Council of 23 July 1996 on the Community guidelines for the development of the trans-European transport networks. *Official Journal of the European Communities* 39, L 228, 9 September 1996, 1-104.

Fujita, M., Krugman, P. Venables, A.J. (1999). *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge, MA: MIT Press.

Fürst, F., Hackl, R., Holl, A., Kramar, H., Schürmann, C., Spiekermann, K., Wegener, M. (1999): *The SASI Model. Model Implementation*. SASI Report D11. Dortmund: Institute of Spatial Planning, University of Dortmund.

Fürst, F., Schürmann, C., Spiekermann, K., Wegener, M. (2000): *The SASI Model. Demonstration Examples*. SASI Report D15. Dortmund: Institute of Spatial Planning, University of Dortmund.

Gálvez, T., Jara-Díaz, S. (1998): On the social valuation of travel time savings. *International Journal of Transport Economics* 25, 205-219.

Hansen, L.P., Heckman, J.J. (1996): The empirical foundations of calibration. *Journal of Economic Perspectives* 10, 87-104.

Hensher, D. (1998): Establishing a fare elasticity regime for urban passenger transport. *Journal of Transport Economics and Policy* 32, 221-246.

Hensher, D.A. (2001): Measurement of the value of time savings. *Journal of Transport Economics and Policy* 35, 71-98.

International Road Union (IRU) (ed.) (2000): *Speed Limits, Goods Transport*. IRU Information Center. www.iru.org/IRU_NEWS/IRU-E.Commerce/Main/Ecommerce.E.asp, October 2000.

IRPUD (2001): *European Transport Networks*. Dortmund: Institute of Spatial Planning, http://irpud.raumplanung.uni-dortmund.de/irpud/pro/ten/ten_e.htm.

Isard, W., Azis, I., Drennan, M., Miller, R., Saltzman, S., Thorbecke, E. (1998): *Methods of Interregional and Regional Analysis*. Brookfield, Vermont: Ashgate Publishing Company.

Jackson, Randall W. (1998): Regionalizing national commodity-by-industry accounts. *Economic Systems Research* 10, 223-38.

Jara-Díaz, S.R. (2000): Allocation and valuation of travel time savings. In: Hensher, D., Button, K (eds.): *Handbooks in Transport, Vol. 1: Transport Modelling*. Oxford: Pergamon Press, 303-319.

Jochimsen, R. (1966): *Theorie der Infrastruktur. Grundlagen der marktwirtschaftlichen Entwicklung*. Tübingen: Mohr.

Jomini, P., Zeitsch, J. F., McDougall, R., Weelsh, A., Brown, S., Hambley, J., Kelly, J. (1991): *SALTER: A General Equilibrium Model of the World Economy, Vol. 1, Model Structure, Database, and Parameters*. Canberra, Australia: Industry Commission.

Kancs D. (2001): Predicting European Enlargement Impacts - A Framework of Interregional General Equilibrium. *Eastern European Economics* 39, 5, 31-63.

Keeble, D., Owens, P.L.; Thompson, C. (1982): 'Regional accessibility and economic potential in the European Community'. *Regional Studies* 16, 419-432.

Keeble, D., Offord, J., Walker, S. (1988): *Peripheral Regions in a Community of Twelve Member States*. Luxembourg: Office for Official Publications of the European Communities.

Krugman, P. (1991): *Geography and Trade*. Leuven: Leuven University Press.

Krugmann, P., Venables, A. (1995): 'Globalisation and the inequality of nations'. *Quarterly Journal of Economics* 60, 4, 857-880.

Leitham, S., Downing, J., Martino, A., Fiorello, D. (1999): *European Transport Forecasts for 2020: The STREAMS Model Results*. Cambridge: Marcial Echenique & Partners Ltd.

- Leontief, W. (1966): *Input-Output-Economics*. Oxford: Oxford University Press.
- Linneker, B. (1997): *Transport Infrastructure and Regional Economic Development in Europe: A Review of Theoretical and Methodological Approaches*. TRP 133. Sheffield: Department of Town and Regional Planning, University of Sheffield.
- Mackie, P., Jara-Díaz, S.R., Fowkes, T. (2001): The value of travel time savings in evaluation. *Transportation Research E* 37, 91-106.
- Marcial Echenique & Partners Ltd. (1998): *The EUNET Project: Socio-Economic and Spatial Impacts of Transport Improvements*. <http://fpiv.meap.co.uk/EUNET.doc>
- Michelin (1992a): *Michelin Reisekarte Großbritannien: Wales, West Country, Midlands*. Stand 1992. Paris: Michelin.
- Michelin (1992b): *Michelin Reisekarte Großbritannien: South East, Midlands, East Anglia*. Stand 1992. Paris: Michelin.
- Miller, R.E., Blair, P.D. (1985): *Input-Output Analysis: Foundations and Extensions*. Englewood Cliffs, NJ: Prentice-Hall.
- Mkmetric (1998): *Subset of the Air Traffic Database*. Karlsruhe: Mkmetric.
- OEEI (2000): *Appraisal of Infrastructural Projects: Guide for Cost-Benefit Analysis*.
- Owen, T., Waters, W., Yong J. (1992): Concepts of price elasticities of transport demand and recent empirical estimates: An interpretive survey. *Journal of Transport Economics and Policy* 26, 139-154.
- Peschel, K. (1981): 'On the impact of geographical distance on interregional patterns of production and trade'. *Environment and Planning A*, 13, 605-269.
- Pfähler, W., Hofmann, U., Bönte, W. (1996): Does extra public infrastructure capital matter? *Finanzarchiv* 53, 68-112.
- Reise- und Verkehrsverlag (1987): *Interrent-Straßenatlas*. Berlin, Gütersloh, München, Stuttgart: Reise- und Verkehrsverlag.
- Rietveld, P., Bruinsma, F. (1998): *Is Transport Infrastructure Effective? Transport Infrastructure and Accessibility: Impacts on the Space Economy*. Advances in Spatial Science. Berlin: Springer.
- Robinson, S., Cattaneo, A., El-Said, M. (2001): Updating and estimating a social accounting matrix using cross entropy methods. *Economic Systems Research* 13, 47 – 64.
- Rus, G. (1990): Public transport demand elasticities in Spain. *Journal of Transport Economics and Policy* 24, 189-201.
- SACTRA (1999): *Transport and the Economy*. London: TSO.

Schürmann, C. (1999): *Quality of Life in European Regions: A Multi-Criteria Analysis*. Arbeitspapier 165. Dortmund: Institute of Spatial Planning.

Schürmann, C., Spiekermann, K., Wegener, M. (1997): *Accessibility Indicators*. SASI Report D5. Berichte aus dem Institut für Raumplanung 39. Dortmund: Institute of Spatial Planning.

Schürmann, C. and Talaat, A. (2000a): *Towards a European Peripherality Index. User Manual*. Paper to the European Commission. Dortmund: Institute of Spatial Planning.

Schürmann, C. and Talaat, A. (2000b): *Towards a European Peripherality Index. Final Paper*. Paper to the European Commission. Dortmund: Institute of Spatial Planning.

Shell (1981). *Der große Shell Atlas Deutschland/Europa 81/82*. Hamburg: Shell.

Shell (1992). *Der große Shell Atlas Deutschland/Europa 92/93*. Hamburg: Shell.

Shiells C.R., Stern, R.M., Deardorff, A.V. (1986): Estimates of the elasticities of substitution between imports and home goods for the United States. *Weltwirtschaftliches Archiv*. 122, 497-519.

Sonesson, T. (2001): Inter-urban travel demand elasticities with emphasis on trip generation and destination substitution. *Journal of Transport Economics and Policy* 35, 301-326.

Spiekermann, K., Wegener, M. (1999): Socio-economic and spatial impacts of trans-European transport networks: the SASI project. In: Meersman, H., Van de Voorde, E., Winkelmans, W. (Hg.): *World Transport Research. Volume 4: Transport Policy*. Selected Proceedings of the 8th WCTR. Elsevier, 1999.

TINA Secretariat (1999): *TINA Transport Infrastructure Needs Assessment. Identification of the Network Components for a Future Trans-European Transport Network in Bulgaria, Cyprus, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia and Slovenia*. Final Paper. Vienna: TINA Secretariat.

UN (1995): *Census of Motor Traffic on Main International Traffic Arteries*. New York: UN Economic Commission for Europe (UN/ECE).

Venables, A.J., Gasiorek, M. (1998): *The Welfare Implications of Transport Improvements in the Presence of Market Failures*. SACTRA Working Paper A, Revised 4/1/98, London: London School of Economics..

Wegener, M., Bökemann, D. (1998): *The SASI Model: Model Structure*. SASI Report D8. Berichte aus dem Institut für Raumplanung 40. Dortmund: Institut für Raumplanung, Universität Dortmund. <http://irpud.raumplanung.uni-dortmund.de/irpud/pro/sasi/ber40.pdf>.

Wegener, M., Schürmann, C., Spiekermann, K. (2000a): *The SASI Model: Model Software*. SASI Report D13. Berichte aus dem Institut für Raumplanung 50. Dortmund: Institut für Raumplanung, Universität Dortmund. <http://irpud.raumplanung.uni-dortmund.de/irpud/pro/sasi/ber50.pdf>.

Wegener, M., Eskilinen, H., Fürst, F., Schürmann, C., Spiekermann, K. (2000b): *Geographical Position. Study Programme on European Spatial Planning*. Bonn: Bundesamt für Bauwesen und Raumordnung (forthcoming).

Vickerman, R., Spiekerman, K., Wegener, M. (1999): Accessibility and economic development in Europe. *Regional Studies* 33, 1-15.

Williams, H.C.W.L. (1977): 'On the formation of travel demand models and economic evaluation measures of user benefit'. *Environment and Planning A* 9, 285-344.